BEST AVAILABLE COPY

U.S.S.N.: 10/521,300

Filed: 1/14/05 Case No.: 21126YP

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 19 April 2001 (19.04.2001)

PCT

(10) International Publication Number WO 01/27322 A2

(51) International Patent Classification⁷:

C12Q 1/68

(21) International Application Number: PCT/US00/21048

(22) International Filing Date: 2 August 2000 (02.08.2000)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data: 60/159,221

13 October 1999 (13.10.1999) US

(71) Applicant (for all designated States except US): ELITRA PHARMACEUTICALS, INC. [US/US]; Suite A, 3510 Dunhill Street, San Diego, CA 92121 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): ZYSKIND, Judith [US/US]; 8514 La Jolla Scenic Drive, La Jolla, CA 92047 (US).

(74) Agent: HUNT, Dale, C.; 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US). (81) Designated States (national): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, IP, KE, KG, KP, KR, KR (utility model), KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

22 A2

776/

(54) Title: CHITOBIASE AS A REPORTER ENZYME

(57) Abstract: The present invention relates to reporter gene constructs encoding a cytoplasmic form of chitobiase (N,N'-diacetyl-chitobiase) and methods of using these reporter gene constructs. The use of a cytoplasmic form of chitobiase as a reporter enzyme is generally applicable to the study of gene expression in organisms which do not contain N-acetyl-β-D-glucosaminidases.

CHITOBIASE AS A REPORTER ENZYME

GOVERNMENT INTEREST IN THE INVENTION

Work performed in developing the present invention was supported, in part, by National Science Foundation grant MCB-9507209, NIH MBRS grant GM45765, and NIH NIGMS MARC F31 GM14967-0451. Accordingly, the United States Government may have certain rights in the invention.

Background of the Invention

Field of the Invention

WO 01/27322

5

10

15

20

25

30

35

Reporter enzymes are enzymes whose activities are easily assayed when present inside cells. In order to study the regulation of a gene whose expression is regulated by various environmental and/or cellular factors or influences, a gene encoding a reporter enzyme may be fused to the coding region or to the regulatory region of the regulated gene.

One of the most popular cytoplasmic reporter enzymes for use in bacteria is ß-galactosidase. It is widely used in the art; however, because bacteria such as *Escherichia coli* contain an endogenous ß-galactosidase encoding gene and ß-galactosidase may be present in the cytoplasm of such bacteria, deletions of the LacZ gene, the source of the enzyme, must be introduced into the host cell line prior to its use. One goal of the present invention is to provide an alternative intracellular enzyme for use as a reporter.

This invention relates to genetic constructs and methods of using a cytoplasmic form of the chitobiase enzyme as a reporter. The invention also comprises expression vectors which express the cytoplasmic form of chitobiase. As used herein, all instances of the terminology "chitobiase" refer to a form of chitobiase which is present in the cytoplasm of the cell. Cytoplasmic forms of chitobiase may be generated via genetic engineering or microbial selection techniques.

Summary of the Invention

One embodiment of the present invention is a method for characterizing a promoter comprising providing a construct comprising the promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase, introducing the construct into host cells, and identifying sequences in the promoter which regulate transcription levels. In one aspect of this embodiment, the cytoplasmic form of chitobiase lacks a signal sequence. In another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein which comprises a cytoplasmic form of chitobiase fused to a heterologous polypeptide. In still another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofe, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi , Vibrio parahaemolyticus, and

5

10

15

20

25

30

Vibrio vulnificus. In a further aspect of this embodiment, the method of identifying sequences which are involved in directing transcription comprises mutagenizing the promoter. In another aspect of this embodiment, the method of identifying sequences which are involved in transcription comprises constructing deletions in the promoter.

Another embodiment of the present invention is a method for identifying a regulatory element capable of directing or regulating transcription within a test nucleic acid sequence comprising providing a construct comprising the test nucleic acid sequence operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase, introducing the construct into host cells, and determining the level of chitobiase activity. In one aspect of this embodiment, the cytoplasmic form of chitobiase lacks a signal sequence. In another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, the fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide. In still another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviplaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi , Vibrio parahaemolyticus, and Vibrio vulnificus. In one aspect of this embodiment, the reporter gene construct is introduced transiently. In another aspect of this embodiment, the reporter gene construct is introduced stably. The host cells may be selected from the group consisting of prokaryotic cells and eukaryotic cells. In another aspect of this embodiment, the method further comprises permeabilizing or lysing the host cells. The permeabilizing or lysing step may comprise treating the host cells with toluene. The step of determining the level of chitobiase activity may be selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate. In another embodiment, the step of determining the level of chitiobiase activity may comprise determining the level of p-nitrophenol released from a substrate. In another aspect of this embodiment, the test nucleic acid sequence comprises a portion of genomic DNA. In a further aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing the host cells to a desired set of environmental conditions. In still another aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting the host cells with a compound to be tested for its influence on the level of transription from siad regulartory element.

Another embodiment of the present invention is a method of detecting successful transformation, comprising the steps of introducing a nucleic acid encoding a cytoplasmic form of chitobiase into host cells, and detecting chitobiase expression in the host cells.

-2-

5

10

15

20

25

30

35

Another embodiment of the present invnetion is a fusion protein-reporter gene construct comprising a promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase fused in frame with a nucleic acid encoding a heterologous polypeptide, wherein the heterologous polypeptide is not β -galactosidase or a portion thereof, and wherein the heterologous polypeptide does not contain a signal peptide. In one aspect of this embodiment, the nucleic acid encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi , Vibrio parahaemolyticus, and Vibrio vulnificus. In another aspect of this embodiment, the nucleic acid further comprises a λ site-specific recombination sequence.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pJMF3.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pJMF4.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pDYK9.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pDYK11.

Another embodiment of the present invention is a host cell comprising a fusion protein-reporter gene construct comprising a promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase fused in frame with a nucleic acid encoding a heterologous polypeptide, wherein the heterologous polypeptide is not β -galactosidase or a portion thereof, and wherein the heterologous polypeptide does not contain a signal peptide. In one aspect of this embodiment, the nucleic acid is integrated into a chromosome of the cell. In another aspect of this embodiment, the nucleic acid is transiently expressed in the host cell.

Another embodiment of the present invention is a nucleic acid encoding a cytoplasmic form of chitobiase in which the signal sequence of native chitobiase has been inactivated or deleted. In one aspect of this embodiment, the signal sequence has been mutated to inactivate it.

An isolated or purified polypeptide comprising a cytoplasmic form of chitobiase fused in frame with a heterologous polypeptide, wherein the heterologous polypeptide is not β -galactosidase or a portion thereof and wherein the heterologous polypeptide does not contain a signal peptide.

Another embodiment of the present invention is an isolated or purified polypeptide comprising a cytoplasmic form of chitobiase in which the signal peptide of native chitobiase has been inactivated or deleted. In one aspect of this embodiment, the signal sequence has been mutated to inactivate it.

Another embodiment of the present invention is a method for monitoring the activity of a promoter comprising providing a construct comprising the promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase, introducing the construct into host cells, and determining the level of chitobiase activity. In one aspect of this embodiment, the cytoplasmic form of chitobiase lacks a signal sequence. In another aspect of this

.3.

5

10

15

20

25

30

embodiment, the nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, the fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide. In one aspect of this embodiment, the nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi , Vibrio parahaemolyticus, and Vibrio vulnificus. In another aspect of this embodiment, the reporter gene construct is introduced transiently. In still another aspect of this embodiment, the reporter gene construct is introduced stably. In a further aspect of this embodiment, the host cells are selected from the group consisting of prokaryotic cells and eukaryotic cells. In yet another aspect of this embodiment, the method further comprises permeabilizing or lysing the host cells. For example, the permeabilizing or lysing step may comprise treating the host cells with toluene. In yet another aspect of this embodiment, the step of determining the level of chitobiase activity may be selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate. In still another aspect of this embodiment, the step of determining the level of chitiobiase activity comprises determining the level of p-nitrophenol released from a substrate. In still another aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing the host cells to a desired set of environmental conditions. In a further aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting the host cells with a compound to be tested for its influence on the level of transription from siad regulartory element. For example, the compound may comprise a compound to be tested for activity as a drug.

Brief Description of the Drawings

Figure 1 illustrates plasmids pJMF3 and pJMF4 containing attP in 2 different orientations and the lac promoter with the first 21 amino acids of $lacZ\alpha$ (from pUC19) fused in-frame to the chb gene. The sequence of the fusion region is shown in Figure 2. Restriction sites shown are found once in the plasmid sequences except for Not, which has 2 sites flanking the P15A origin.

Figure 2 illustrates the sequence of the lac promoter and the chitobiase fusion found in pJMF3 and pJMF4 (SEQ ID NOS: 15 and 16). Fusion between $lacZ\alpha$ (from pUC19) and chb [Soto-Gill, R.W. et al, J. Biol. Chem. 264:14778-14782 (1998)] is indicated by (/); start of transcription is indicated by (+1). Sequence and binding sites in the lac promoter regulatory region are found in Dickson, R.C. et al., Science 187:27-32 (1975). Restriction enzyme sites shown in Figure 2 are found once in the plasmid sequences except for Sph1, which has 2 sites; these different

4.

5

10

15

20

25

30

restriction sites can be used to replace the *lac* promoter with another promoter together with part of a coding region to produce an in-frame fusion with *chb*.

Figure 3 depicts integration of *chb* fusions into the chromosome by site-specific recombination between *attB* and *attP*. The steps involved are described below and in Diederich, L.L, et al., Plasmid 28:14-24 (1992).

Figure 4 illustrates plasmids pDYK9 and pDYK11 which contain *dnaA-chb* fusions. pDYK9 is deleted for the *rpmH* regulatory region. The orientation of *attP* in pDYK9 and pDYK11 is the same as that of pJMF3. After integration at *attB* of the larger *Not*1 fragment, the *dnaA* promoters are oriented to transcribe in the same direction as replication fork movement.

Figure 5 illustrates the *rmpH-dnaA* regulatory region and *dnaA*-chitobiase fusion (SEQ ID NO.: 17). The *dnaA* box and promoters are shaded, and the coding region of the *dnaA* gene is striped. The fusion contains two amino acids (between the backslashes) from pUC19. The region cloned into pDYK9 is between *SphI* primer II and *KpnI* primer, and the region cloned into pDYK11 is between *SphI* primer I and *KpnI* primer. The numbers above the primers refer to the nucleotide in the sequence amplified. For sequence numbering and the locations of promoters and protein binding sites, see Froelich, J.M. et al., J. Bacteriol. 178:6006-6012 (1996); Hansen, F.G. et al., EMBO J. 1:1043-1048 (1982) and Messer, W. and C.W. Weigel, Initiation of Chromosome Replication. p. 1579-1601 (1996) In F.C. Neidhart, R. Curtiss III, J.L. Ingraham, E.CC. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter and H.E. Umbarger (Eds.), Escherichia coli and Salmonella Cellular and Molecular Biology., ASM Press, Washington, D.C.

Detailed Description of the Preferred Embodiment

Reporter genes and reporter gene constructs play a number of important roles in a variety of molecular biology techniques. For example, reporter genes may be used to determine whether a sequence contains a promoter or other cis-acting element which directs transcription, such as an enhancer. In addition, reporter genes may be used to identify regulatory sites in promoters or other cis-acting elements and to determine the effects of mutating these regulatory sites on the level of gene expression directed by the promoters or other cis-acting elements. Reporter genes may also be used to detect successful transformation. In addition, reporter genes may be used to monitor gene expression under various conditions and to identify drugs.

The structure of a reporter gene construct containing the cytoplasmic form of chitobiase will vary according to its purposes. When the reporter gene construct is a vector, one must decide between a vector that incorporates the reporter gene into the host's genome or one that replicates extrachromosomally, such as a plasmid. When the reporter construct is not designed to integrate into the host genome, the vector can contain an origin of replication with activity in the host cell of interest. This feature provides the reporter gene vector the ability to replicate within the host cell in which it has been introduced. In addition to the origin of replication, reporter constructs often contain a promoter, a multiple cloning site, a selectable marker, and of course a reporter gene. Reporter constructs for use in eukaryotic cells may also contain a polyA site adjacent to the reporter gene.

-5-

Given the utility of reporter gene constructs, it is not surprising that a number of cytoplasmic reporter gene constructs and different reporter genes are available for use by those of skill in the art. For example, the cytoplasmic reporter enzymes chloramphenical acetyltransferase (CAT), firefly luciferase, β-glucuronidase (GUS), green fluorescent protein (GFP), and β-galactosidase have been used extensively. However, such reporters all have individual shortcomings that may limit or preclude their usage under some conditions. For example, high levels of GFP are toxic to the cell. In addition, reporter enzymes are not expressed equally in all cell types nor are they equally stable when expressed in all cell types. Furthermore, there is a recognized need for multiple reporter enzymes that can be assayed independently of one another in order to simultaneously study the regulation of multiple genes within a single cell type. Therefore, there exists a continuing need to identify reporter enzymes with useful properties.

5

10

15

20

25

30

The cytoplasmic enzyme B-galactosidase is widely used as a reporter gene in various microbiological and molecular biological studies. This enzyme is used in both *in vitro* and *in vivo* assays. The wide acceptance of this reporter system results, in part, because it is non-isotopic and extremely flexible. It is used in a number of assay formats and has an extremely broad linear range. Nevertheless, because B-galactosidase is present in the cytoplasm of various host cells such as *Escherichia coli*, deletion of the *lacZ* gene, the source of the enzyme, is often required prior to its use in a host cell system. One goal of the present invention was to provide an alternative intracellular enzyme for use as a reporter.

An extensive discussion of various molecular biology techniques is available in Ausubel, et al., (eds) "Short Protocols in Molecular Biology," Wiley and Sons, Inc., New York (1997). Examples of such techniques include isolating and preparing DNA for manipulation, gel electrophoresis, polymerase chain reaction (PCR), determining nucleic acid sequences, screening nucleic acid libraries, mutagenesis of DNA, and introducing DNA into host cells.

The structure of a reporter gene construct will vary according to its purposes. The reporter gene constructs are constructed according to standard techniques of molecular biology well known in the art.

When the reporter gene construct is a vector, one must decide between a vector that incorporates the reporter gene into the host's genome or one that replicates extrachromosomally, such as a plasmid. When integration of the reporter gene is a desired result, a reporter gene construct will contain sequences that will facilitate incorporation.

One example of integration sequences that can be included in a reporter gene construct is the λ attP site. This site permits a single copy of the reporter gene to be incorporated into a host bacterial genome. Integration-promoting sequences with utility in mammalian cells include the long terminal repeats found in retroviral genomes. These sequences promote viral genome integration in a host genome and have been used extensively by those of skill in the art to promote the integration of exogenous sequences in mammalian host cells.

When the reporter construct is not designed to integrate into the host genome, it is common that the vector contain an origin of replication with activity in the host cell of interest. This feature provides the reporter gene vector the ability to replicate within the host cell in which it has been introduced.

5

10

15

20

25

30

35

In addition to the origin of replication, reporter constructs will often contain additional features that promote the expression of the nucleic acid sequence or sequences contained in the construct. These additional sequences can include a polyA site, a multiple cloning site, a drug resistance marker, and of course a reporter gene.

The present invention relates to the use of chitobiase as a reporter gene. The chitobiase may be used as a reporter in bacteria, plants, mammalian cells and other host cell lines. One possible alternative to using 8-galactosidase as a reporter gene was to develop cytoplasmic N,N'-diacetylchitobiase W-acetyl-8-D-glucosaminidase, EC 3.2.1.30) for use as a reporter enzyme. One advantage of the enzyme N-acetyl-8-D-glucosaminidase or "chitobiase" over 8-galactosidase is that genes encoding chitobiase are missing from many bacteria, including E. coli, some fungi, and some eukaryotic cells. Thus, it is not necessary to engineer many host cells to lack reporter activity as is the case with 8-galactosidase. The present invention also relates to various protein expression vectors that can be used to express the reporter gene. In addition, the present invention may be used in conjunction with other reporter enzymes in systems in which the regulation or activities of multiple genes is to be studied simultaneously.

Chitobiase is one of two enzymes that hydrolyze chitin, an abundant insoluble polysaccharide, to its monomeric unit, N-acetylglucosamine (GlcNac). Chitobiase is known to be present in a number of organisms. For example, the chitobiase enzyme is known to be present in various genera including Arabidopsis, Bacillus, Bombyx, Bos, Caenorhabditis, Candida, Dictyostelium, Entamoeba, Felis, Homo, Korat, Lactobacillus, Leishmania, Mus, Pisum, Porphyromonas, Pseudoalteromonas, Rattus, Serratia, Streptomyces, Sus, Trichoderma, and Vibrio. Specific examples of organisms known to contain chitobiase include Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi, Vibrio parahaemolyticus, and Vibrio vulnificus.

One source of the enzyme is the marine bacterium, Vibrio harveyi. Escherichia coli cells harboring a plasmid carrying the chb gene from Vibrio harveyi were reported to produce the enzyme, which was found to be associated with the outer membrane of the bacterial cells. These are described in Jannatipour, M. et al., "Translocation of Vibrio harveyi N,N'-diacetylchitobiase to the outer membrane of Escherichia coli.," J. Bacteriol. 169:3785-3791 (1987) and Soto-Gil & Zyskind, N,N'-Diacetylchitobiase of Vibrio harveyi primary structure, processing, and evolutionary relationships. J. Biol Chem. 264:14778-14782 (1989).

The present invention contemplates expressed cytoplasmic forms of chitobiase in various forms. In one embodiment, the signal sequence is deleted from the amino terminal portion of the protein. Presumably the removal of this sequence results in the expression of a cytoplasmic form of the enzyme that is not secreted from the host cell or incorporated into the membrane of the host cell producing the enzyme.

The present invention also contemplates the generation of fusion proteins comprising a fusion polypeptide joined in frame to chitobiase. Preferably, the fusion polypeptide comprises a polypeptide other than chitobiase, such

.7.

as a heterologous protein. The heterologous polypeptide may comprise a polypeptide having a biological activity (such as an enzymatic or other activity besides activity as an immunogen) or the heterologous polypeptide may not have a biological activity. The heterologous polypeptide does not include a signal sequence which directs its secretion. Preferably the heterologous polypeptide is not β -galactosidase or a portion thereof. Thus, the fusion reporter gene construct contains a sequence encoding the fusion polypeptide genetically fused in frame with a sequence encoding chitobiase. In one embodiment, this fusion may remove the amino-terminal signal peptide sequence of chitobiase and replace it with a heterologous protein.

5

10

15

20

25

30

In another embodiment, the fusion protein construct comprises a chitobiase gene sequence that has been truncated to remove at least the signal peptide sequence of the gene. Alternatively, mutations may be introduced into the signal peptide sequence so that it is no longer functional. Such mutations may be introduced using a variety of techniques familiar to those skilled in the art, including site directed mutagenesis, cassette mutagenesis, and chemical mutagenesis.

Once the reporter gene construct is made it is introduced into a host cell line for testing. Host cells of prokaryotic and eukaryotic origin can be used with the reporter gene constructs of the present invention. A variety of methods are available to introduce reporter gene constructs into host prokaryotic cells. For example, bacteria can be transformed using calcium chloride, electroporation, or viral vectors such as the filamentous phages. These and other prokaryotic transformation protocols are well known in the art.

Alternatively, the sequence encoding chitobiase may be introduced in eukaryotic cells, including yeast, mammalian, plant, and insect cells. For example, the sequence encoding chitobiase may be inserted into a yeast artificial chromosome, a yeast plasmid, a bovine papilloma virus vector or other extrachromosomal vector, a retroviral vector, a Ti-plasmid, or a baculovirus vector. A variety of such vectors are known to those skilled in the art. The vectors may be introduced into any of the yeast, mammalian, plant, and insect cells familiar to those skilled in the art.

The introduction of the reporter gene construct into mammalian cells can likewise utilize a number of transfection protocols well known to those of skill in the art. As discussed above, transfections can be transient or stable. Examples of suitable transfer protocols include calcium phosphate transfection, DEAE-Dextran, electroporation, liposome-mediated transfection, and viral transfection. These and other eukaryotic transformation protocols are well known in the art.

Following introduction of the reporter gene construct into the host cell of interest, the enzymatic activity of the reporter gene is measured. Preferably, the chitobiase assays are performed after permeabilizing or lysing the host cells. There are a variety of cell permeabilization and cell lysis procedures available to those of ordinary skill in the art, including methods such as sonication or lysozyme treatment. One embodiment of the present invention uses toluene treatment to permeabilize cells. The details of this method are discussed in D.Y. Kalabat et al., BioTechniques 25:1030-1035 (1998) and Miller, J.H. A Short Course in Bacterial Genetics, CSH Laboratory Press, Cold Spring Harbor, NY 1992.

-8-

5

10

15

20

25

30

35

Cellular chitobiase activity can be measured quantitatively by following the hydrolysis of chitobiase substrates. Examples of substrates with utility in chitobiase activity assays include N,N'-diacetylchitobiase (chitobiase), p-nitrophenyl-N-acetyl-8-D-glucosaminide (PNAG)(Sigma Chemical, St. Louis, MO), and 5-bromo-4-chloro-3-indolyl-N-acetyl-B-D-glucosaminide (X-Gluc)(Sigma Chemical, St. Louis, MO). Other substrates are also contemplated for use in the assays of the present invention.

Products produced by the hydrolysis of the chitobiase substrates are monitored using various means familiar to those skilled in the art. For example, various optical means are known to those skilled in the art. One such optical means may comprise detection of chemiluminescent or fluorescent products released from a substrate. Alternatively, the level of chitobiase activity may be determined by measuring the amount of light absorbed by a product produced from a substrate or measuring a decrease in the amount of a detectable substrate. In one embodiment, *p*-nitrophenol is released from the substrate and measured at 400 nm. Other monitoring methods well known in the art can be used to quantitate signals produced in the chitobiase assay. These may inleade use of radioactive substrates or substrates having radiofrequency tags. In another embodiment, blue/white colony indicator plates are used to monitor enzyme activity.

Another embodiment of the present invention is a kit. One aspect of this embodiment includes a reporter gene construct comprising a vector containing a chitobiase reporter gene. The reporter gene construct also contains a multiple cloning site containing a variety of restriction endonuclease cutting sites that facilitate the introduction of exogenous DNA into the construct.

The kit embodiment of the present invention also includes those components necessary to assay for chitobiase activity produced by the reporter gene construct. For example, in one embodiment, the kit will include a supply of a suitable chitobiase substrate whose metabolism into product by the reporter enzyme can be assayed.

EXAMPLES

The following Examples are disclosed to assist in the understanding of the present invention. The Examples below should not be construed to limit the scope of the invention and such variations of the invention now known or later developed, which would be within the purview of one of ordinary skill in the art and are considered to fall within the scope of the invention hereinafter claimed.

Example 1

Construction of Vectors

Vectors were constructed using polymerase chain reaction (PCR) products that were cloned first at the *Smal* or *Eco*RV restriction sites of plasmid pBluescript II (pKSII+)(Stratagene; San Diego, CA). The nucleotide sequence of all PCR products was determined using standard techniques well known in the art. Unmethylated plasmid DNA was isolated from an *E. coli dam* strain when cutting with the *Bcl*I enzyme was required.

The pDYK9 (SEQ ID NO: 11) plasmid was constructed by ligating a *Sphl-Kpnl* PCR product containing the *dnaA* promoter region into plasmid pRSG196 that contains the *V. harveyi* chitobiase gene (see Jannatipour, M., et al.). Briefly, the *V. harveyi* chitobiase gene was prepared by cloning a *Sau3A* partial digest of *V. harveyi* DNA into the

.g.

5

10

15

20

25

30

35

single BamHI site with the tet gene of pMK2004 [See Soto-Gill & Zyskind, "Cloning of Vibrio harveyi chitinase and chitobiase genes in Escherichia coli," pp. 209-223, In J.P. Zikakis (ed.), Chitin, chitosan, and related enzymes," Academic Press, Inc., New York (1984)]. The clones containing the gene of interest were detected by the presence of the yellow p-nitrophenylate product after individual colonies of the clone bank were sprayed with 10 mM PNAG dissolved in 100 mM sodium phosphate, pH 7.0. Restriction maps of the positive clones were made according to standard techniques well known in the art.

The *chb* game was subcloned in a 3.5-kb *Eco*Rl fragment from pRSG14 [See Jannatipour, M. et al., "Translocation of *Vibrio harveyi N,N'*-diacetylchitobiase to the outer membrane of *Escherichia coli.*," J. Bacteriol. 169:3785-3791 (1987)] into the *Eco*Rl site of pUC19 [Yanisch-Perron, c.J. et al., Gene 33: 103-119 (1985)]. Plasmid pRSG196 was constructed by deleting a 0.5-kb *Sst*l fragment from one of these clones [See Jannatipour, M. et al., "Translocation of *Vibrio harveyi N,N'*-diacetylchitobiase to the outer membrane of *Escherichia coli.*," J. Bacteriol. 169:3785-3791 (1987)].

The pRSG196 plasmid was also cut with *Sph*I and *Kpn*I to accommodate the insertion of the *dnaA* fragment. Plasmid pAC17 (described in Chiaramello & Zyskind, "Coupling of DNA replication to growth rate in *Escherichia coli*: a possible role for guanosine tetraphosphate," J. Bacteriol. 1*72*:2013-2019 (1992)] served as template with primers 5'-GCA CAT GCA TGC TGG TCA TTA AAT TTT CC-3' (SEQ ID NO 1) and 5'-CGG GGT ACC AAC TCA TCC TGC AAT CG-3' (SEQ ID NO 2) producing a PCR product 374 bp long that contains 353 bp from the *dnaA* promoter region (bases 583 to 935, numbering according to Hansen, F. G., et al., "The nucleotide sequence of the *dnaA* gene promoter and of the adjacent *rpmH* gene, coding for the ribosomal protein L34, of *Escherichia coli*," EMBO J. 1:1043-1048 (1982)]. The forward primer (*Sph*I primer II) contained an *Sph*I site and the reverse primer (*Kpn*I primer) contained a *Kpn*I site for cloning. This created an in-frame fusion between the amino-terminal 17 amino acids of DnaA and the carboxy-terminal end of chitobiase deleted for the amino-terminal 22 amino acids including the signal peptide.

The next step in constructing the pDYK9 vector was taking the 3270 bp *Dral- Hind*III (partial digest) fragment containing the *dnaA-chb* fusion and ligating it to the pACYC184 2555 bp *Hinc*II-*Hind*III fragment carrying chloramphenicol-resistance (Cm^r) and the P15A origin. This fragment was described in Chang & Cohen, "Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid," J. Bacteriol. 134:1141-1156 (1987).

A Not! site was introduced at the Acc! site after digestion with Acc!, treatment with Mung bean nuclease, and ligation to phosphorylated Not! linkers (New England Biolabs, Inc.; Beverly, MA). A Not! site was introduced at the Ase! site after digestion with Ase!, treatment with Mung bean nuclease, and ligation to phosphorylated Not! linkers (New England Biolabs, Inc.; Beverly, MA). An Xbal-Sph! PCR product containing the rmBt1t2 terminator was ligated into this plasmid cut with the same enzymes creating pDYK7. E. coli chromosomal DNA served as template with primers 5'-CTA GTC TAG ATG CCG AAC TCA GAA GTG A-3' (SEQ ID NO 3) and 5'-GCA CAT GCA TGC GGG GGA TGG CTT GTA GAT-3' (SEQ ID NO 4) to produce a PCR product 357 bp long that contains bases 6534 to 6869 from the rmB operon [numbering according to Brosius, J., et al., "Gene organization and primary structure of a

5

10

15

20

25

30

35

ribosomal RNA operon from *Escherichia coli.*" J. Mol. Biol. 148:107-127 (1981)] and includes the complex transcription termination region of this operon which is described in Orosz, A., et al., "Analysis of the complex transcription termination region of the *Escherichia coli rmB* gene," Eur. J. Biochem. *201*:653-659 (1991). The forward primer contained an *Xba*l site and the reverse primer contained an *Sph*l site.

A Bch-Smal PCR product containing the λ attP site was ligated into pDYK7 cut with Tth111, treated with Mung bean nuclease, then digested with Bch. Plasmid pHN894 [described in Goodman, S. D., et al., "Deformation of DNA during site-specific recombination of bacteriophage lambda: Replacement of IHF protein by HU protein or sequence-directed bends," Proc. Natl. Acad. USA 89:11910-11914 (1992)] served as template with the forward primer 5'-CAT GAT CAT GCG ACA GGT TTG ATG A-3' (SEQ ID NO 5) and the reverse primer 5'-GGG GGC GCC TAC CTT TCA CGA G-3' (SEQ ID NO 6) producing a PCR product 466 bp long that contains the λ attP site. The 466 bp PCR product containing the λ attP site was first cloned into the Smal site of the pKSII+ plasmid to produce pDYK8. The forward primer contains a Bch site and the reverse primer contains G's at the 5' end in order to recreate a Smal site when cloning the PCR fragment into a Smal site. This PCR product includes bases -211 to +241 from the center of the attP core and the sequence required for optimum λ attP site integration as described in Nagaraja & Weisberg, "Specificity determinants in the attachment sites of bacteriophages HK022 and λ ," J. Bacteriol. 172:6540-6550 (1990). The orientation of attP is such that when the fusion is integrated at attB, the transcription direction of dnaAp1 and p2 promoters is the same as replication fork movement mimicking the orientation at the wild type dnaA promoters.

Vector pDYK11 (SEQ ID NO: 12) was constructed by ligating an Sphl-Kpnl PCR product containing the rpmH-dnaA promoter region into pDYK9 also digested with Sphl and Kpnl. Plasmid pAC17, described in Chiaramello & Zyskind, (1992), served as template with the primers 5'-CAT GCA TGC ATG AAA CGA TGG ACA CC-3' (SEQ ID NO 7) and 5'-CGG GGT ACC AAC TCA TCC TGC AAT CG-3' (SEQ ID NO 8) to produce a PCR product 616 bp long that contains 598 bp from the rpmH-dnaA regulatory region [bases 338 to 935, numbering according to Hansen, F. G., et al., (1982)]. The forward primer (Sphl primer I) contained an Sphl site and the reverse primer (Kpnl primer) contained a Kpnl site for cloning. SEQ ID NO: 18 provides the complete coding sequence of the DNA encoding the dnaA/chitobiase fusion protein. The complete sequence of the dnaA/chitobiase fusion protein is provided in SEQ ID NO: 19.

The vector pJMF3 (SEQ ID NO: 13) was constructed by first ligating an *Ase*I linker (5'-CATTAATGCATG-3' (SEQ ID NO 9) self-hybridized) into the *Sph*I site of pDYK11. The pUC19, [described in Yanisch-Perron C., et al., "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors," Gene *33*:103-19 (1985)] *AseI-Kpn*I fragment containing the *lacPO*-polylinker region was ligated to this plasmid after digestion with *Ase*I and *Kpn*I. The resulting in-frame fusion between the amino-terminal 21 amino acids of the pUC19 *lacZ*(α) peptide [Yanisch-Perron C., et al., 1985] and the carboxy-terminal end of chitobiase deleted for the amino-terminal 22 amino acids is identical to the protein fusion in pRSG196 [Jannatipour, M., et al., (1995)].

Vector pJMF4 (SEQ ID NO: 14) was constructed by ligating the BamHi-EcoRV fragment of pDYK8 containing the λ attP site into pDYK7 which had been digested with Tth1111, treated with Mung bean nuclease, then

digested with *BcI*, to create pTKP9. This reversed the orientation of the *attP* site relative to the *attP* site in pDYK9. The pDYK11 *BspMII-KpnI* fragment containing the *rpmH-dnaA* promoter region was then ligated into pTKP9 digested with the same enzymes to create pTKP11. This reversed the orientation of the *attP* site relative to the *attP* site in pDYK11. An *Asel* linker [5'-CATTAATGCATG-3' (SEQ ID NO 10) self-hybridized) was ligated into the *SphI* site of pTKP11 to create pJMF2. The pUC19 [as described in Yanisch-Perron C., et al., (1985)] *Asel-KpnI* fragment containing the *IacPO*-polylinker region was ligated to pJMF2 cut with *Asel* and *KpnI* to create pJMF4.

Example 2

Site-specific Recombination

To move the chitobiase fusions in pDYK9 and pDYK11 to the attB site in the chromosome, Not1 fragments from these plasmids were self-ligated and transformed or electroporated into strain WM2269 (DH5 α containing pLDR8) [See Zyskind, J. W. and S. I. Bernstein (1992) "Recombinant DNA Laboratory Manual," Academic Press, San Diego, CA]. Plasmid pLDR8, described in Diederich, L. L. J., et al., "New cloning vectors for integration into the λ attachment site attB of Escherichia coli chromosome," Plasmid 28:14-24 (1992), expresses integrase from the λ P_R promoter and contains the λ cl_{857} repressor gene, a kanamycin resistance gene, and a temperature-sensitive origin of replication. The transformed or electroporated cells were incubated at 42°C with shaking for 30 min then moved to 37°C for 1 h followed by selection on Luria broth agar plates containing chloramphenicol (25 μ g/ml) at 42°C. Transformants were screened for loss of kanamycin resistance and, therefore, loss of pLDR8.

Example 3

Bacteriophage P1 Transduction

Transduction with P1 bacteriophage by the method of Zyskind & Bernstein, Recombinant DNA Laboratory Manual. Academic Press, San Diego, CA (1992) was used to construct strains and to confirm the chromosomal location of the *dnaA-chb* fusions. Cotransduction of Cm^r (carried by the fusion) and *galK* (linked to *attB*) indicated that Cm^r and *galK* are linked on the chromosomes of strains DYK9W, DYK9F, DYK11W, and DYK11F.

Example 4

25 p-Nitrophenol Chitobiase Activity Assay

5

10

15

20

30

35

Chitobiase activity is located in the cytoplasm when its signal peptide is replaced by fusion with another peptide. Accordingly, chitobiase assays are performed on permeabilized or lysed cells. A variety of permabilization or cell lysis protocols are available to liberate the enzyme from within a cell population being tested. One such protocol involves toluene-treated cells washed once with M9 salts according to the method of Miller, J.H. A Short Course in Bacterial Genetics, CSH Laboratory Press, Cold Spring Harbor, NY 1992. The toluenized cells are placed in a chitobiase buffer (10 mM Tris-HCl, pH 8.0, and 0.5 M NaCl) and 666 µM PNAG, the chitobiase enzyme substrate. NaCl is included because chitobiase has approximately 80% of full activity in the absence of salt, with maximal activity occurring between 0.25 and 0.6 M NaCl. Toluenized cells (0.772 ml) are preincubated at 28°C and the reaction started with the addition of 0.228 ml PNAG (1 mg/ml). After incubation at 28°C, the reaction is stopped by the addition of 1 ml of 1 M Tris base. The release of p-nitrophenol is measured at 400 nm and turbidity at 550 nm. p-

Nitrophenol release is measured immediately at 400 nm with a molar absorptivity of 1.8 x 10^3 liters mol⁻¹ cm⁻¹. Units are calculated after subtracting the light scattering factor (1.5 x $0D_{550}$) from $0D_{400}$ of the sample. The normalizing factor of 1.5 was determined previously by measuring the light scattering ratio of bacteria at $0D_{400}$ and $0D_{550}$. One unit of chitobiase activity is the amount of enzyme that catalyzes the formation of 1 pmol of p-nitrophenol per min at 28°C. For comparison to Miller units of 8-galactosidase (described in Miller, J. H., A Short Course in Bacterial Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1992)], the units are normalized to 1 ml of culture at $0D_{450}$ – 1.

Example 5

Determination of Chitobiase Activity in Cells Containing Vectors pJM3 and pJM4

(Figure 1) contain the *lacPO* promoter with the first 21 amino acids of $lacZ\alpha$ (from pUC19) fused in-frame to the *chb* gene. These plasmids also contain the λ phage attP recombination site in different orientations, the gene encoding chloramphenical acetyltransferase (*cat*), and a ribosomal terminator, *rmBt1t2*, inserted upstream of the *lac-chb* fusion to prevent read-through from other promoters. The chitobiase activity associated with these plasmids (Table 1) is high in the absence of IPTG because of titration of *lac* repressor expressed from a single copy chromosomal gene. Induction by IPTG is approximately 10-fold (Table 1).

Table 1. Chitobiase Activity of lacZ-chb Fusion*

Plasmid	Chitobiase ^b (U ^c)		
	-IPTG	+IPTG (1 mM)	
pJMF3 in DH5α	668 ± 45	9320 ± 347	
pJMF4 in DH5α	788 ± 44	7188 ± 477	

^{*} Overnight cultures were diluted 1:1000 into 50 mL prewarmed LB and grown to $0D_{450}$ = 0.1. 1 mM IPTG was added to half of the culture, and growth continued to $0D_{450}$ = 0.3.

20

25

5

10

15

The *lac* promoter can be replaced with another promoter and a fusion protein created with chitobiase by cutting with *Sph*I or *Ase*I and either *Sal*I, *Kpn*I, or *Sst*I (Figure 2). Fusions created with these vectors can be moved to the chromosome by site specific recombination at the λ attB site to permit single copy analysis of the activity of the promoter. The protocol, as described more fully in Diederich, et al., "New cloning vectors for integration into the λ attachment site attB of *Escherichia coli* chromosome," Plasmid 28:14-24 (1992), involves two components, (i) a circular DNA containing the λ attachment site, attP, the promoter-chb gene fusion, and the cat gene, and (ii) a helper plasmid, pLDR8, which contains the *int* gene under the control of the temperature-sensitive repressor, cl857 and a temperature-sensitive origin of replication. The plasmid is digested with *Not*I to remove the P15A origin, and the

⁶ Triplicate samples were assayed. Mean chitobiase activities are given with standard deviations.

One unit of chitobiase activity is 1 pmol of ρ -nitrophenol/min at 28°C. Units given for 1 mL of culture at OD_{450} = 1.

fragment containing the chitobiase fusion is self-ligated prior to transformation into cells containing pLDR8. Integration occurs by site-specific recombination between *attP* and *attB* (17.4 min on the *E. coli* chromosome, Figure 3).

Example 6

Use of the Chitobiase Reporter Enzyme to Study dnaA Gene regulation

Two plasmids, pDYK9 and pDYK11, discussed in the Examples above, were constructed to assess the regulation of the *dnaA* gene using chitobiase as a reporter enzyme. These plasmids differ by the absence of the *rpmH* regulatory region in pDYK9 (Figure 4). These fusions were moved from the plasmid to the chromosomal *attB* site for single copy analysis as described above. After transformation of strain WM2269 with the ligated DNA, integration occurred by site specific recombination between the *attP* and the *attB* sites. The orientation of *attP* in pDYK9 and pDYK11 is such that when the fusion is integrated at *attB*, the transcription direction of *dnaAp1* and *dnaAp2* promoters is the same as movement of the replication fork. This orientation is the same as at the *dnaA* wild type location. The genetic location of the fusions was confirmed in the Cm^f transformants by demonstrating cotransduction of Cm^f and *galK*.

The fusions created in strain WM2269 were moved by P1 transduction to MG1655, creating strain DYK9W with pDYK9 and strain DYK11W with pDYK11. Deletion of the *rpmH* promoters had very little effect (1.4-fold) on chitobiase activity (Table 2, compare lines 1 and 3)

Table 2. Chitobiase and β -Gal Activities of dnaA-chb and dnaA-lacZ Fusions in WT and fis Mutant Backgrounds

Strains		β- Gal ^ε	
	Chitobiase* (U*)	(Miller U ⁴)	
DYK11W fis+	30.0 ± 2.1		
DYK11F <i>fis</i> ::985	80.0 ± 1.2		
DYK9W fis+	44.0 ± 2.1		
DYK9F <i>fis</i> ::985	96.0 ± 3.5		
RB220 <i>fis</i> +		59.8 ± 7.9	
TP220 <i>fis</i> ::767		115.2 ± 7.2	

Triplicate samples were assayed during exponential growth. Mean chitobiase activities are given with standard deviation.

5

10

15

20

One unit of chitobiase activity is 1 pmol of ρ -nitrophenol per min at 28°C. Units given for 1 mL of culture at $OD_{450} = 1$.

^c Data from Reference 8.

Unit defined in Reference 13.

Fis protein binds to a site in the *dnaAp2* promoter that covers the -35 sequence (Figure 5), and appears to be a repressor of DnaA expression. A fusion protein with B-galactosidase activity that is expressed from the *rpmH-dnaA* regulatory region has increased B-galactosidase activity (1.9-fold) in a *fis** mutant when compared to Fis wild type cells [as described in Froelich, J. M., et al., "Fis binding in the *dnaA* operon promoter region," J. Bacteriol. *178*:6006-6012 (1996), data shown in Table 3]. Similarly, the absence of Fis leads to a greater than 2-fold increase in chitobiase activity of the DnaA-chitobiase fusion protein for the DYK9F and DYK11F strains, comparable in extent to that observed with the *dnaA-lacZ* fusion strain, TP220 (Table 3).

In the reporter gene constructs discussed in the Example, all upstream transcriptional activity was prevented from entering the *chb* reporter gene. The plasmid vectors, pJMF3 and pJMF4, described in the Examples above, contained the *rmbt1t2* terminator upstream of the promoter fusion, which prevented readthrough from chromosomal promoters near the insertion site. Only chitobiase activity originating from the promoters of interest was expressed.

The attP site in these vectors allowed integration at the chromosomal attB in a specific orientation, depending on the vector used. With these vectors, any chitobiase fusion involving an essential gene can be moved to the chromosome, thus permitting single copy analysis with a chromosomal orientation similar to the wild-type gene.

Example 7

5

10

15

20

25

30

Identification of Promoters in Test Sequences

A nucleic acid prospectively containing a promoter is inserted upstream of a nucleic acid encoding a cytoplasmic form of chitobiase as described above. For example, the nucleic acid prospectively containing a promoter may be inserted into a restriction site in a sequence containing a plurality of restriction sites, such as a polylinker, which is located upstream of the nucleic acid encoding chitobiase. The test sequence may comprise any nucleic acid to be tested for promoter activity. In one embodiment, the test sequence may comprise a genomic DNA sequence. For example, the genomic DNA sequence may be a randomly generated DNA fragment, such as a fragment generated using shotgun cloning techniques, a restriction fragment, or any other sequence.

The vectors containing the test sequence upstream of the nucleic acid encoding chitobiase are introduced into an appropriate host cell. The level of chitobiase activity is assayed and compared to the level obtained from a control vector which lacks an insert in the cloning site. The presence of an elevated expression level in cell containing the vector containing the insert with respect to the level in cells containing the control vector without the insert indicates the presence of a promoter in the insert.

In some embodiments, the activity of the promoter in the test sequence may be assayed after exposure of the host cells to conditions which may influence the level of transcription from the promoter. For example, the environment of the host cells may be altered to determine whether the transcription level is influenced by environmental factors, including factors such as temperature, pH, nutrients, or availability of oxygen. In such analyses, chitobiase levels are assayed under a variety of environmental conditions to determine the effects of the environmental conditions on transcription levels from the promoter. In addition, the activity of the promoters may be examined in the presence or absence of compounds to be

5

10

15

20

25

30

tested for regulatory activity. For example, the activity of the promoters may be tested by determining the levels of chitobiase produced in the presence or absence of compounds to be tested for activity as drugs.

Promoter sequences within the test sequences may be further defined by constructing nested deletions in the test sequences using conventional techniques such as Exonuclease III digestion. The resulting deletion fragments can be inserted into the promoter reporter vector to determine whether the deletion has reduced or obliterated promoter activity as determined by measuring chitobiase activity in cells containing the deletion vectors. In this way, the boundaries of the promoters may be defined. If desired, potential individual regulatory sites within the promoter may be identified using techniques such as site directed mutagenesis or linker scanning to obliterate potential transcription factor binding sites within the promoter individually or in combination. The effects of these mutations on transcription levels may be determined by inserting the mutations into the cloning sites in the promoter reporter vectors and measuring the levels of chitobiase produced from the mutated promoters.

The activity of known promoters may also be monitored by operably linking them to a nucleic acid encoding a cytoplasmic form of chitobiase. The activity of the promoters may be analyzed under various environmental conditions as described above. In addition, the activity of the promoters may be analyzed in the presence or absence of compounds to be tested for the ability to affect transcription from the promoters. For example, the compounds may be tested for activity as drugs.

In some embodiments, the chitobiase reporter constructs may be used in systems for identifying compounds that modulate cell surface protein-mediated activity or compounds which modulate the activities of intracellular signaling systems. Techniques for using reporter genes to identify compounds which modulate cell surface protein-mediated activity have been described in U.S. Patent Number 5,401,629 and U.S. Patent Number 5,436,128. Briefly, in such methods, a construct comprising a promoter operably linked to a nucleic acid encoding a reporter enzyme is introduced into cells which express the cell surface protein and cells which do not express the cell surface protein. Each of the cells are contacted with test compounds and the effects of these compounds on transcription levels is measured by determining the level of activity of the reporter enzyme. The level of expression of the reporter gene in cells expressing the cell surface protein is compared to the level in cells which do not express the cell surface protein to identify compounds that modulate cell surface protein activity.

Similarly, the chitobiase reporter constructs may be used to identify compounds which influence the activity of intracellular signaling pathways, such as cAMP-based or phosphorylation-based pathways. In such methods, a promoter which is activated via such pathways is operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase. The cells are contacted with test compounds. Those compounds which activate the pathway to which the promoter responds will produce an enhanced level of chitobiase activity in the cells as compared to the level of chitobiase activity in control cells which have not been contacted with the test compound.

Example 8

Detecting Successful Transformation or Transfection Using Chitobiase

A vector comprising a sequence encoding a cytoplasmic form of chitobiase operably linked to a sequence capable of directing transcription of the chitobiase gene is introduced into a host cell. The host cells are contacted with a chitobiase substrate and those host cells which contain chitobiase activity are identified as cells which were successfully transformed or transfected. In some embodiments, a portion or replica of a colony may be lysed or permeabilized prior and the lysate or permeabilized cells may be contacted with the chitobiase substrate.

CONCLUSION

New gene reporter systems that use chitobiase have been described. Chitobiase has advantages over other reporter gene systems in that chitobiase is not found in many cell lines traditionally used in reporter gene systems.

Finally, the forgoing examples are not intended to limit the scope of the present invention, which is set forth in the following claims. In particular, various equivalents and substitutions will be recognized by those of ordinary skill in the art in view of the foregoing disclosure, and these are contemplated to be within the scope of the present invention.

15

10

5

WHAT IS CLAIMED IS:

1. A method for characterizing a promoter comprising:

providing a construct comprising said promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase;

5

introducing the construct into host cells; and identifying sequences in said promoter which regulate transcription levels.

- 2. The method of Claim 1, wherein said cytoplasmic form of chitobiase lacks a signal sequence.
- 3. The method of Claim 2, wherein said nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, said fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide.

10

15

20

- 4. The method of Claim 1, wherein said nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi, Vibrio parahaemolyticus, and Vibrio vulnificus.
- 5. The method of Claim 1, wherein said method of identifying sequences which are involved in directing transcription comprises mutagenizing said promoter.
- 6. The method of Claim 1, wherein said method of identifying sequences which are involved in transcription comprises constructing deletions in said promoter.
- 7. A method for identifying a regulatory element capable of directing or regulating transcription within a test nucleic acid sequence comprising:

providing a construct comprising said test nucleic acid sequence operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase;

25

introducing said construct into host cells; and determining the level of chitobiase activity.

- 8. The method of Claim 7, wherein said cytoplasmic form of chitobiase lacks a signal sequence.
- 9. The method of Claim 8, wherein said nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, said fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide.

30

10. The method of Claim 7, wherein said nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9. Rattus norvegicus, Serratia

marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi , Vibrio parahaemolyticus, and Vibrio vulnificus.

- 11. The method of Claim 7, wherein said reporter gene construct is introduced transiently.
- 12. The method of Claim 7, wherein said reporter gene construct is introduced stably.

5

10

15

20

25

30

35

- 13. The method of Claim 7, wherein said host cells are selected from the group consisting of prokaryotic cells and eukaryotic cells.
 - 14. The method of Claim 7, further comprising permeabilizing or lysing said host cells.
- 15. The method of Claim 14, wherein said permeabilizing or lysing step comprises treating said host cells with toluene.
- 16. The method of Claim 7, wherein said step of determining the level of chitobiase activity is selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate.
- 17. The method of Claim 7, wherein said step of determining the level of chitiobiase activity comprises determining the level of ρ -nitrophenol released from a substrate.
 - 18. The method of Claim 7, wherein said test nucleic acid sequence comprises a portion of genomic DNA.
- 19. The method of Claim 7, wherein said step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing said host cells to a desired set of environmental conditions.
- 20. The method of Claim 7, wherein said step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting said host cells with a compound to be tested for its influence on the level of transription from siad regulartory element.
 - 21. A method of detecting successful transformation, comprising the steps of: introducing a nucleic acid encoding a cytoplasmic form of chitobiase into host cells; and detecting chitobiase expression in said host cells.
- 22. A fusion protein-reporter gene construct comprising a promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase fused in frame with a nucleic acid encoding a heterologous polypeptide, wherein said heterologous polypeptide is not -galactosidase or a portion thereof and wherein said heterologous polypeptide does not contain a signal peptide.
- 23. The nucleic acid of Claim 22, wherein said nucleic acid encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi, Vibrio parahaemolyticus, and Vibrio vulnificus.

24. The nucleic acid of Claim 22, further comprising a λ site-specific recombination sequence.

- 25. A reporter gene construct comprising plasmid pJMF3.
- 26. A reporter gene construct comprising plasmid pJMF4.
- 27. A reporter gene construct comprising plasmid pDYK9.
- 28. A reporter gene construct comprising plasmid pDYK11.
- 29. A host cell comprising the construct of Claim 22.

5

15

20

25

30

- 30. The host cell of Claim 29 wherein said nucleic acid is integrated into a chromosome of said cell.
- 31. The host cell of Claim 29, wherein said nucleic acid is transiently expressed in said host cell.
- 32. A nucleic acid encoding a cytoplasmic form of chitobiase in which the signal sequence of native chitobiase has been inactivated or deleted.
 - 33. The nucleic acid of Claim 32, wherein the signal sequence has been mutated to inactivate it.
 - 34. An isolated or purified polypeptide comprising a cytoplasmic form of chitobiase fused in frame with a heterologous polypeptide, wherein said heterologous polypeptide is not -galactosidase or a portion thereof and wherein said heterologous polypeptide does not contain a signal peptide.
 - 35. An isolated or purified polypeptide comprising a cytoplasmic form of chitobiase in which the signal peptide of native chitobiase has been inactivated or deleted.
 - 36. The polypeptide of Claim 35, wherein the signal sequence has been mutated to inactivate it.
 - 37. A method for monitoring the activity of a promoter comprising:

providing a construct comprising said promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase;

introducing said construct into host cells; and determining the level of chitobiase activity.

- 38. The method of Claim 37, wherein said cytoplasmic form of chitobiase lacks a signal sequence.
- 39. The method of Claim 38, wherein said nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, said fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide.
- 40. The method of Claim 37, wherein said nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of Alteromonas sp. 0-7, Arabidopsis thaliana, Bacillus subtilis, Bombyx mori, Bos taurus, Caenorhabditis elegans, Candida albicans, Dictyostelium discoideum, Entamoeba histolytica, Felis catus, Homo sapiens, Korat cats, Lactobacillus casei, Leishmania donovani, Mus musculus, Pisum sativum, Porphyromonas gingivalis, Pseudoalteromonas sp. S9, Rattus norvegicus, Serratia marcescens, Streptomyces plicatus, Streptomyces thermoviolaceus, Sus scrofa, Trichoderma harzianum, Vibrio furnissii, Vibrio harveyi, Vibrio parahaemolyticus, and Vibrio vulnificus.
 - 41. The method of Claim 37, wherein said reporter gene construct is introduced transiently.
 - 42. The method of Claim 37, wherein said reporter gene construct is introduced stably.

43. The method of Claim 37, wherein said host cells are selected from the group consisting of prokaryotic cells and eukaryotic cells.

- 44. The method of Claim 37, further comprising permeabilizing or lysing said host cells.
- 45. The method of Claim 44, wherein said permeabilizing or lysing step comprises treating said host cells with toluene.
 - 46. The method of Claim 37, wherein the step of determining the level of chitobiase activity is selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate.
 - 47. The method of Claim 37, wherein said step of determining the level of chitiobiase activity comprises determining the level of ρ -nitrophenol released from a substrate.

10

15

- 48. The method of Claim 37, wherein said step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing said host cells to a desired set of environmental conditions.
- 49. The method of Claim 37, wherein said step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting said host cells with a compound to be tested for its influence on the level of transription from siad regulartory element.
- 50. The method of Claim 49, wherein said compound comprises a compound to be tested for activity as a drug.

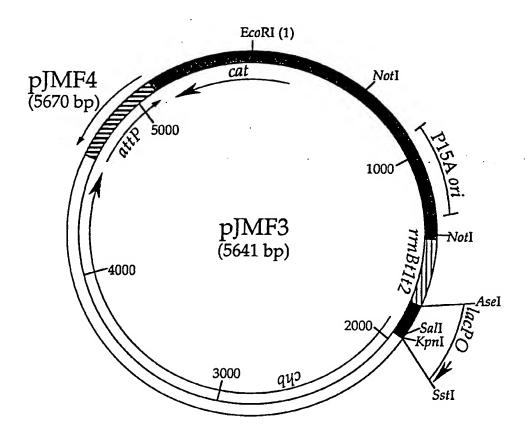


Figure 1

SEQ ID NO: 15:

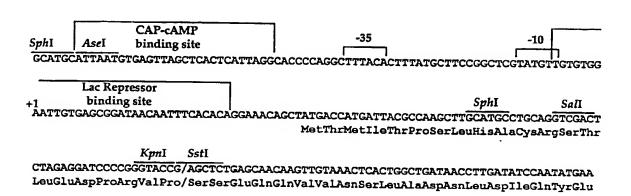


Figure 2

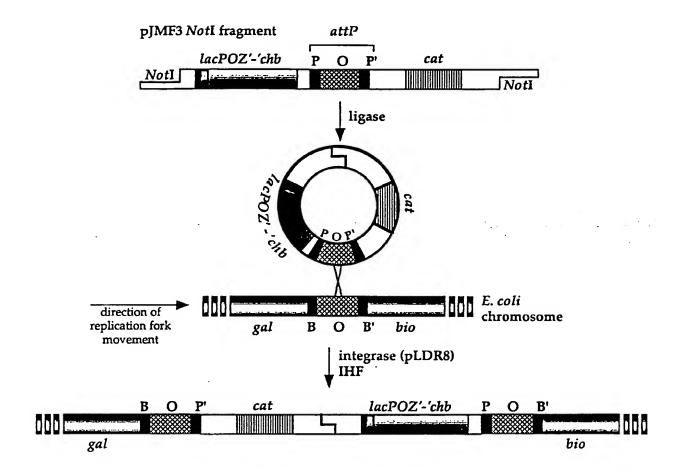


Figure 3

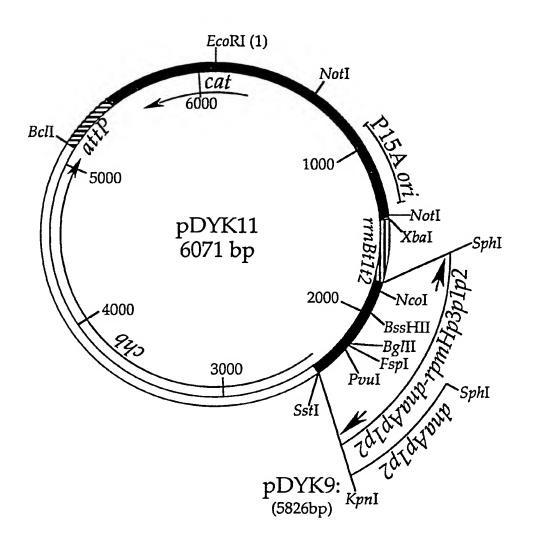
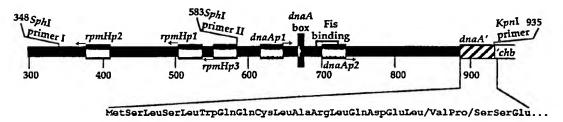


Figure 4



SEQ ID NO: 17

Figure 5

SEQUENCE LISTING

<110>	ELITRA PHARMACEUTICALS, INC. Zyskind, Judith W.	
<120>	CHITOBIASE AS A REPORTER ENZYME	
<130>	ELITRA.006VPC	
	60/159,221 1999-10-13	
<160>	19	
<170>	FastSEQ for Windows Version 4.0	
<210> <211> <212> <213>	29	
<220> <223>	Oligonucleotide	
<400> gcacat	1 :gcat gctggtcatt aaattttcc	29
<210> <211> <212> <213>	26	
<220> <223>	Oligonucleotide	
<400> cggggt	2 acca actcatcctg caatcg	26
<210> <211> <212> <213>	28	
<220> <223>	Oligonucleotide	
<400> ctagto	3 staga tgccgaactc agaagtga	28
<210> <211> <212> <213>	30	
<220> <223>	Oligonucleotide	
<400> gcacat <210>	gcat gcgggggatg gcttgtagat	30

PCT/US00/21048 WO 01/27322 <211> 25 <212> DNA <213> Artificial Sequence <223> Oligonucleotide <400> 5 catgatcatg cgacaggttt gatga 25 <210> 6 <211> 22 <212> DNA <213> Artificial Sequence <220> <223> Oligonucleotide <400> 6 22 gggggcgcct acctttcacg ag <210> 7 <211> 26 <212> DNA <213> Artificial Sequence <220> <223> Oligonucleotide <400> 7 catgcatgca tgaaacgatg gacacc 26 <210> 8 <211> 26 <212> DNA <213> Artificial Sequence <220> <223> Oligonucleotide cggggtacca actcatcctg caatcg 26 <210> 9 <211> 12 <212> DNA <213> Artificial Sequence <220> <223> Oligonucleotide <400> 9 cattaatgca tg 12 <210> 10 <211> 12 <212> DNA <213> Artificial Sequence <220>

<223> Oligonucleotide

```
<400> 10
cattaatgca tg
                                                                        12
<210> 11
<211> 5826
<212> DNA
<213> Artificial Sequence
<220>
<223> pDYK9
<400> 11
qaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg gataaaactt
                                                                       60
qtqcttattt ttctttacqq tctttaaaaa ggccgtaata tccagctgaa cggtctggtt
                                                                       120
ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat gccattggga
                                                                      180
tatatcaacg gtggtatatc cagtgatttt tttctccatt ttagcttcct tagctcctga
                                                                      240
aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat ggtgaaagtt
                                                                      300
qqaacctctt acgtqccqat caacgtctca ttttcgccaa aagttggccc agggcttccc
                                                                      360
ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaqqtat
                                                                      420
ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt
                                                                      480
qtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt tcagctactg
                                                                      540
acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtgcggc
                                                                      600
cqcactqqct tactatqttq qcactqatga gggtgtcagt gaagtgcttc atgtgqcagg
                                                                      660
agaaaaaagg ctgcaccggt gcgtcagcag aatatgtgat acaggatata ttccgcttcc
                                                                      720
tegeteactg actegetacg eteggtegtt egactgegge gageggaaat ggettacgaa
                                                                      780
cqqqqcqqaq atttcctgga agatgccagg aagatactta acagggaagt gagagggccg
                                                                      840
cggcaaagcc gtttttccat aggctccgcc cccctgacaa gcatcacgaa atctgacgct
                                                                      900
caaatcagtg gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggcgg
                                                                      960
ctccctcgtg cgctctcctg ttcctgcctt tcggtttacc ggtgtcattc cgctgttatg
                                                                     1020
qccqcqtttq tctcattcca cqcctqacac tcaqttccgg gtaggcagtt cgctccaagc
                                                                     1080
tggactgtat gcacgaaccc cccgttcagt ccgaccgctg cgccttatcc ggtaactatc
                                                                     1140
gtettgagte caacceggaa agacatgcaa aagcaccact ggcagcagce actggtaatt
                                                                     1200
                                                                     1260
gatttagagg agttagtctt gaagtcatgc gccggttaag gctaaactga aaggacaagt
                                                                     1320
tttggtgact gcgctcctcc aagccagtta cctcggttca aagagttggt agctcagaga
accttcqaaa aaccqccctq caaggcggtt ttttcgtttt cagagcaaga gattacgcgc
                                                                     1380
agaccaaaac gatctcaaga agatcatctt atgcggccgc atcagataaa atatttctag
                                                                     1440
atgccgaact cagaagtgaa acgccgtagc gccgatggta gtgtggggtc tccccatgcg
                                                                     1500
agagtaggga actgccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt
                                                                     1560
tcqttttatc tqttqtttqt cggtgaacgc tctcctgagt aggacaaatc cgccgggagc
                                                                     1620
qqatttqaac qttqcqaaqc aacqqcccqq aqqqtqqcqq qcaqqacqcc cqccataaac
                                                                     1680
tgccaggcat caaattaagc agaaggccat cctgacggat ggcctttttg cgtttctaca
                                                                     1740
                                                                     1800
aactetteet qteqteatat etacaageea teeceegeat getggteatt aaatttteea
atatgcggcg taaatcgtgc ccgcctcgcg gcaggatcgt ttacacttag cgagttctgg
                                                                     1860
aaagtcctgt ggataaatcg ggaaaatctg tgagaaacag aagatctctt gcgcagttta
                                                                     1920
ggctatgatc cgcggtcccg atcgttttgc aggatcttga tcgggcatat aaccgcagac
                                                                     1980
ageggttegt gegteaccet caageagggt ettttegacg taegteaaca ateatgaatg
                                                                     2040
tttcagcctt agtcattatc gacttttgtt cgagtggagt ccgccgtgtc actttcgctt
                                                                     2100
tggcagcagt gtcttgcccg attgcaggat gagttggtac cgagctctga gcaacaagtt
                                                                     2160
gtaaactcac tggctgataa ccttgatatc caatatgaag tgttaactaa ccatggtgct
                                                                     2220
aacqaaqqtc ttqcqtqcca aqatatqqqc qcaqaatqqq cttcttqtaa caaaqtaaac
                                                                     2280
atgacqcttq ttaaccaaqq tqaaqctqtt gactcaaaag attgggctat ttacttccac
                                                                     2340
agcattcgtc tgattctgga tgttgacaac gagcagttca aaatctctcg tgtaacgggt
                                                                     2400
gacctacata agctagaacc aacagataag tttgacggct tcgctgccgg tgaagaggtt
                                                                     2460
gttcttccat tggttggtga atactggcaa ctatttgaaa ctgacttcat gccgggtgca
                                                                     2520
ttcgtttctg ctccaaacgc agaacctaag atgattgctt ctctaaatac tgaagatgtt
                                                                     2580
gcgtcttttg tgacgggtct tgaaggtaac aacctaaaac gtacaccaga tgacaacaat
                                                                     2640
gtatttgcaa acgctgtgtc tcgttttgag aaaaacgaag acctagcaac acaagacgta
                                                                     2700
tcaaccacgt tactaccaac accaatgcac gttgaagcgg gtaaaggcaa agtagatatc
                                                                     2760
qcqqatqqta ttqcqctqcc taaaqacqca ttcqatqcqa ctcaqttcqc agcqattcaa
                                                                     2820
qatcqtqcaq aaqtqqtaqq tqtqqacqtt cqtqqtqatc ttcctqtaaq catcactqtt
                                                                     2880
```

```
gttcctgcag acttcaccgg tgaattagca aaatctggtg cttacgaaat gagcatcaaa
                                                                      2940
ggcgacggta ttgtgattaa agcgttcgac caagcaggcg ctttctacgc agtacaatct
                                                                      3000
atctttggcc tggtagatag ccaaaatgct gattctctac cacaactgtc tattaaagat
                                                                      3060
gcgcctcgtt ttgattaccg tggtgtgatg gtggatgtgg ctcgtaactt ccactctaag
                                                                      3120
qacqcaatcc ttgcaacqct agaccaaatg gcagcqtaca agatgaacaa acttcacctt
                                                                      3180
cacctaaccg atgatgaagg ctggcgttta gaaatcccgg gtctgcctga gctgacagaa
                                                                      3240
gtgggtgcta accgttgttt cgatacacaa gagaaaagct gtttactgcc tcagcttggc
                                                                      3300
totgqtccaa cqacagacaa ctttggctct ggctacttca gcaaagcaga ctacgtggaa
                                                                      3360
atcttgaaat acgcgaaagc acgtaacatt gaagtgattc cagaaatcga tatgccagct
                                                                      3420
cacqctcqtq caqcagtagt atcaatggaa gctcgttacg accgcctaat ggaagaaggt
                                                                      3480
aaaqaaqctq aaqcqaacqa ataccqtctg atggatcctc aagatacatc aaacqtaacq
                                                                      3540
acggttcagt tctacaataa gcaaagcttc atcaacccat gtatggaatc ttcaactcgc
                                                                      3600
tttgttgata aggtgatttc agaagtggca gcaatgcacc aagaagctgg cgctccacta
                                                                      3660
acaacttggc acttcggtgg tgacgaagcg aagaacatca agctaggtgc tggtttccaa
                                                                      3720
qacqttaacq cagaagataa agtaagctgg aaaggcacga ttgacctgtc taaacaaqac
                                                                      3780
aagccqtttq cacaqtctcc acaatgtcag acgctaatca cagatggcac agtcagtgac
                                                                      3840
tttgctcacc taccaagcca cttcgcggaa gaagtgtcga agattgttgc tgagaaaggc
                                                                      3900
attccaaact tccaagcttg gcaagatggt ttgaaataca gtgacggcga aaaagcgttc
                                                                      3960
qctacaqaaa atactcgcgt aaacttctgg gacgttctgt actggggcgg tacttcctca
                                                                      4020
gtgtacgagt ggtctaagaa aggttacgac gtgattgttt ctaacccaga ttacgtgtac
                                                                      4080
atggatatgc catacgaagt tgacccgaaa gagcgtggtt actactgggc aacacgtgca
                                                                      4140
acggatactc gtaagatgtt tggctttgca ccagagaaca tgcctcaaaa cgcagaaact
                                                                      4200
tctgtagatc gcgatggcaa tggctttact ggtaaaggtg aaatcgaagc gaaacctttc
                                                                      4260
tacqqtctat ctqcacaact ttqqtctqaq acaqtacqta acqacqaqca atacqaqtac
                                                                      4320
atggtattcc ctcqcqtcct cqctqctqct cagcqtqcat gqcaccqtqc tqactqqqaa
                                                                      4380
aacgactaca aagttggtgt tgagtactcg caaaactcta atctagttga taaagcatcg
                                                                      4440
ctaaaccaag actacaaccg ctttgcgaac gtacttggtc aacgtgaact ggctaagcta
                                                                      4500
gaaaaatcag gtattgacta ccgcctacca gtaccaggtg caaaagtaga agatggtaag
                                                                      4560
ctagcaatga acgttcagtt ccctggcgta acgcttcaat actctctgga tggtgagaac
                                                                      4620
tggttgactt atgcagacaa cgctcgtcca aatgtaactg gtgaagtctt catccgctcg
                                                                      4680
gtatctgcga caggtgagaa ggtaagccgt atcactagcg tgaaataata gcgctcagta
                                                                      4740
ttcactaaaa tcataqttcc ttactcaaaq ccctcaactt atqttqqqqq ctttqtttat
                                                                      4800
ttttcttcqq aaaataaqcq tqatcatqcq acaqqtttqa tqacaaaaaa ttaqcqcaaq
                                                                      4860
aagacaaaaa tcaccttqcq ctaatgctct qttacaggtc actaatacca tctaagtagt
                                                                      4920
tgattcatag tgactgcata tgttgtgttt tacagtatta tgtagtctgt tttttatgca
                                                                      4980
anatctaatt taatatattg atatttatat cattttacgt ttctcgttca gcttttttat
                                                                      5040
                                                                      5100
actaagttgg cattataaaa aagcattgct tatcaatttg ttgcaacgaa caggtcacta
tcagtcaaaa taaaatcatt atttgatttc aattttgtcc cactccctgc ctctgtcatc
                                                                      5160
acgatactgt gatgccatgg tgtccgactt atgcccgaga agatgttgag caaacttatc
                                                                      5220
gettatetge tteteataga gtettgeaga caaactgege aactegtgaa aggtaggege
                                                                      5280
ccccgqtcqa atttqctttc gaatttctgc cattcatccg cttattatca cttattcagg
                                                                      5340
cgtagcacca ggcgtttaag ggcaccaata actgccttaa aaaaattacg ccccgccctg
                                                                      5400
ccactcatcg cagtactgtt gtaattcatt aagcattctg ccgacatgga agccatcaca
                                                                      5460
gacggcatga tgaacctgaa tcgccagcgg catcagcacc ttgtcgcctt gcgtataata
                                                                      5520
tttgcccatg gtgaaaacgg gggcgaagaa gttgtccata ttggccacgt ttaaatcaaa
                                                                      5580
actggtgaaa ctcacccagg gattggctga gacgaaaaac atattctcaa taaacccttt
                                                                      5640
agggaaatag gccaggtttt caccgtaaca cgccacatct tgcgaatata tgtgtagaaa
                                                                      5700
ctgccggaaa tcgtcgtggt attcactcca gagcgatgaa aacgtttcag tttgctcatg
                                                                      5760
gaaaacggtg taacaagggt gaacactatc ccatatcacc agetcaccgt etttcattge
                                                                      5820
                                                                      5826
catacq
<210> 12
<211> 6071
<212> DNA
<213> Artificial Sequence
<220>
<223> pDYK11
<400> 12
                                                                        60
gaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg gataaaactt
```

gtgcttattt	ttctttacgg	tctttaaaaa	ggccgtaata	tccagctgaa	cggtctggtt	120
ataggtacat	tgagcaactg	actgaaatgc	ctcaaaatqt	tctttacgat	gccattggga	180
	gtggtatatc					240
	aactcaaaaa					300
	acgtgccgat					360
	gggacaccag					420
ggtattaata	gggacaccag	gacccaccca	taccacatta	ctastttsct	ctataggtat	480
	caaagtgcgt					
	tgctccagtg					540
	gcgtaacggc					600
	tactatgttg					660
	ctgcaccggt					720
tcgctcactg	actcgctacg	ctcggtcgtt	cgactgcggc	gagcggaaat	ggcttacgaa	780
cggggcggag	atttcctgga	agatgccagg	aagatactta	acagggaagt	gagagggccg	840
	gtttttccat					900
	gtggcgaaac					960
	cgctctcctg					1020
	tctcattcca					1080
	gcacgaaccc					1140
						1200
	caacccggaa					
	agttagtctt					1260
	gcgctcctcc					1320
	aaccgccctg					1380
	gatctcaaga					1440
atgccgaact	cagaagtgaa	acgccgtagc	gccgatggta	gtgtggggtc	tccccatgcg	1500
agagtaggga	actgccaggc	atcaaataaa	acgaaaggct	cagtcgaaag	actgggcctt	1560
	tgttgtttgt					1620
	gttgcgaagc					1680
	caaattaagc					1740
	gtcgtcatat					1800
	tgattaaaga					1860
						1920
	taccgcgctt					1980
	agcatggcgc					
	ggatctttat					2040
	cggcgtaaat					2100
	cctgtggata					2160
	tgatccgcgg					2220
cagacagcgg	ttcgtgcgtc	accctcaagc	agggtctttt	cgacgtacgt	caacaatcat	2280
gaatgtttca	gccttagtca	ttatcgactt	ttgttcgagt	ggagtccgcc	gtgtcacttt	2340
cgctttggca	gcagtgtctt	gcccgattgc	aggatgagtt	ggtaccgagc	tctgagcaac	2400
	ctcactggct					2460
	aggtcttgcg					2520
	gcttgttaac					2580
	tcgtctgatt					2640
	acataagcta					2700
aggregater	tccattggtt	gatazztact	accasetatt	tgaaactgac	ttcataccaa	2760
						2820
	ttctgctcca					
	ttttgtgacg					2880
	tgcaaacgct					2940
	cacgttacta					3000
	tggtattgcg					3060
ttcaagatcg	tgcagaagtg	gtaggtgtgg	acgttcgtgg	tgatcttcct	gtaagcatca	3120
ctgttgttcc	tgcagacttc	accggtgaat	tagcaaaatc	tggtgcttac	gaaatgagca	3180
	cggtattgtg					3240
	tggcctggta					3300
	tcgttttgat					3360
	aatccttgca					3420
_	aaccgatgat					3480
						3540
	tgctaaccgt					
	tccaacgaca					3600
tggaaatctt	gaaatacgcg	aaagcacgta	acattgaagt	gattccagaa	accgatatgc	3660
cagctcacgc	tcgtgcagca	gtagtatcaa	tggaagctcg	ttacgaccgc	ctaatggaag	3720

```
aaggtaaaga agctgaagcg aacgaatacc gtctgatgga tcctcaagat acatcaaacg
                                                                      3780
taacqacggt tcagttctac aataagcaaa gcttcatcaa cccatgtatg gaatcttcaa
                                                                      3840
ctcgctttgt tgataaggtg atttcagaag tggcagcaat gcaccaagaa gctggcgctc
                                                                      3900
cactaacaac ttggcacttc ggtggtgacg aagcgaagaa catcaagcta ggtgctggtt
                                                                      3960
tecaagaegt taaegeagaa gataaagtaa getggaaagg caegattgae etgtetaaae
                                                                      4020
aaqacaagcc gtttgcacag tctccacaat gtcagacgct aatcacagat ggcacagtca
                                                                      4080
gtgactttgc tcacctacca agccacttcg cggaagaagt gtcgaagatt gttgctgaga
                                                                      4140
aaggcattcc aaacttccaa gcttggcaag atggtttgaa atacagtgac ggcgaaaaag
                                                                      4200
cqttcqctac agaaaatact cqcqtaaact tctqqqacqt tctqtactqq qqcqqtactt
                                                                      4260
cctcagtgta cgagtggtct aagaaaggtt acgacgtgat tgtttctaac ccagattacq
                                                                      4320
tqtacatqqa tatqccatac qaaqttqacc cgaaaqaqcq tqqttactac tqqqcaacac
                                                                      4380
qtqcaacqga tactcqtaag atgtttgqct ttqcaccaga gaacatqcct caaaacqcaq
                                                                      4440
aaacttctgt agatcgcgat ggcaatggct ttactggtaa aggtgaaatc gaagcgaaac
                                                                      4500
ctttctacqq tctatctqca caactttgqt ctgagacagt acqtaacqac qaqcaatacq
                                                                      4560
aqtacatggt attecetege gteetegetg etgeteageg tgeatggeae egtgetgaet
                                                                      4620
gggaaaacga ctacaaagtt ggtgttgagt actcgcaaaa ctctaatcta gttgataaag
                                                                      4680
categetaaa ccaagactac aacegetttg cgaacgtact tggtcaacgt gaactggcta
                                                                      4740 -
agctagaaaa atcaggtatt gactaccgcc taccagtacc aggtgcaaaa gtagaagatg
                                                                      4800
gtaagctagc aatgaacgtt cagttccctg gcgtaacgct tcaatactct ctggatggtg
                                                                      4860
agaactggtt gacttatgca gacaacgctc gtccaaatgt aactggtgaa gtcttcatcc
                                                                      4920
qctcqqtatc tqcqacaggt gagaaggtaa gccgtatcac tagcgtgaaa taatagcqct
                                                                      4980
caqtattcac taaaatcata gttccttact caaagccctc aacttatgtt gggggctttg
                                                                      5040
tttatttttc ttcqqaaaat aaqcqtqatc atqcqacaqq tttqatqaca aaaaattaqc
                                                                      5100
qcaaqaaqac aaaaatcacc ttgcgctaat gctctgttac aggtcactaa taccatctaa
                                                                      5160
qtaqttqatt cataqtqact gcatatgttg tgttttacag tattatgtag tctgttttt
                                                                      5220
atgcaaaatc taatttaata tattgatatt tatatcattt tacgtttctc gttcagcttt
                                                                      5280
tttatactaa gttggcatta taaaaaagca ttgcttatca atttgttgca acgaacaggt
                                                                      5340
cactatcagt caaaataaaa tcattatttg atttcaattt tgtcccactc cctgcctctg
                                                                      5400
tcatcacgat actgtgatgc catggtgtcc gacttatgcc cgagaagatg ttgagcaaac
                                                                      5460
ttatcgctta tctgcttctc atagagtctt gcagacaaac tqcgcaactc gtgaaaggta
                                                                      5520
ggcgcccccg gtcgaatttg ctttcgaatt tctgccattc atccqcttat tatcacttat
                                                                      5580
tcaggcgtag caccaggcgt ttaagggcac caataactgc cttaaaaaaa ttacgccccg
                                                                      5640
ccctqccact catcqcaqta ctqttqtaat tcattaaqca ttctqccqac atqqaaqcca
                                                                      5700
tcacagacgg catgatgaac ctgaatcgcc agcggcatca gcaccttgtc gccttgcgta
                                                                      5760
taatatttgc ccatggtgaa aacgggggcg aagaagttgt ccatattggc cacgtttaaa
                                                                      5820
tcaaaactgg tgaaactcac ccagggattg gctgagacga aaaacatatt ctcaataaac
                                                                      5880
cctttaggga aataggccag gttttcaccg taacacgcca catcttgcga atatatgtgt
                                                                      5940
agaaactgcc ggaaatcgtc gtggtattca ctccagagcg atgaaaacgt ttcagtttgc
                                                                      6000
tcatggaaaa cggtgtaaca agggtgaaca ctatcccata tcaccagctc accgtctttc
                                                                      6060
attqccatac q
                                                                      6071
<210> 13
<211> 5641
<212> DNA
<213> Artificial Sequence
<220>
<223> pJMF3
<400> 13
                                                                        60
gaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg gataaaactt
gtgcttattt ttctttacgg tctttaaaaa ggccgtaata tccagctgaa cggtctggtt
                                                                       120
ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat gccattggga
                                                                       180
tatatcaacg gtggtatatc cagtgatttt tttctccatt ttaqcttcct taqctcctqa
                                                                       240
aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat ggtgaaagtt
                                                                       300
ggaacctctt acgtgccqat caacgtctca ttttcgccaa aagttggccc agggcttccc
                                                                       360
ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat
                                                                       420
ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt
                                                                       480
gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt tcagctactg
                                                                       540
                                                                       600
acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtgcggc
cgcactggct tactatgttg gcactgatga gggtgtcagt gaagtgcttc atgtggcagg
                                                                       660
                                                                       720
agaaaaaagg ctgcaccggt gcgtcagcag aatatgtgat acaggatata ttccgcttcc
```

	actcgctacg					780
	atttcctgga					840
cggcaaagcc	gtttttccat	aggctccgcc	cccctgacaa	gcatcacgaa	atctgacgct	900
caaatcagtg	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	cccctggcgg	960
	cgctctcctg					1020
	tctcattcca					1080
	gcacgaaccc					1140
	caacccggaa					1200
	agttagtctt					1260
tttggtgact	gcgctcctcc	aagccagtta	cctcggttca	aagagttggt	agctcagaga	1320
	aaccgccctg					1380
agaccaaaac	gatctcaaga	agatcatctt	atgcggccgc	atcagataaa	atatttctag	1440
atgccgaact	cagaagtgaa	acgccgtagc	gccgatggta	gtgtggggtc	tccccatgcg	1500
agagtaggga	actgccaggc	atcaaataaa	acgaaaggct	cagtcgaaag	actgggcctt	1560
	tgttgtttgt					1620
ggatttgaac	gttgcgaagc	aacggcccgg	agggtggcgg	gcaggacgcc	cgccataaac	1680
tgccaggcat	caaattaagc	agaaggccat	cctgacggat	ggcctttttg	cgtttctaca	1740
	gtcgtcatat					1800
	caccccaggc					1860
gtgagcggat	aacaatttca	cacaggaaac	agctatgacc	atgattacgc	caagcttgca	1920
tgcctgcagg	tcgactctag	aggatccccg	ggtaccgagc	tctgagcaac	aagttgtaaa	1980
	gataaccttg					2040
	tgccaagata					2100
gcttgttaac	caaggtgaag	ctgttgactc	aaaagattgg	gctatttact	tccacagcat	2160
	ctggatgttg					2220
	gaaccaacag					2280
	ggtgaatact					2340
ttctgctcca	aacgcagaac	ctaagatgat	tgcttctcta	aatactgaag	atgttgcgtc	2400
ttttgtgacg	ggtcttgaag	gtaacaacct	aaaacgtaca	ccagatgaca	acaatgtatt	2460
	gtgtctcgtt					2520
cacgttacta	ccaacaccaa	tgcacgttga	agcgggtaaa	ggcaaagtag	atatcgcgga	2580
	ctgcctaaag					2640
	gtaggtgtgg					2700
	accggtgaat					2760
	attaaagcgt					2820
	gatagccaaa					2880
	taccgtggtg					2940
	acgctagacc					3000
	gaaggctggc					3060
	tgtttcgata					3120
	gacaactttg					3180
	aaagcacgta					3240
	gtagtatcaa					3300
	aacgaatacc					3360
	aataagcaaa					3420
	atttcagaag					3480
	ggtggtgacg					3540
	gataaagtaa					3600
	tctccacaat					3660
	agccacttcg					3720
	gcttggcaag					3780
	cgcgtaaact					3840
	aagaaaggtt					3900
	gaagttgacc					3960
	atgtttggct					4020
	ggcaatggct					4080
	caactttggt					4140
	gtcctcgctg					4200
	ggtgttgagt					4260
	aaccgctttg					4320
atcaggtatt	gactaccgcc	taccagtacc	aggtgcaaaa	gtagaagatg	gtaagctagc	4380

```
aatgaacgtt cagttccctg gcgtaacgct tcaatactct ctggatggtg agaactggtt
                                                                      4440
                                                                      4500
qacttatgca gacaacgctc gtccaaatgt aactggtgaa gtcttcatcc gctcggtatc
tgcgacaggt gagaaggtaa gccgtatcac tagcgtgaaa taatagcgct cagtattcac
                                                                      4560
taaaatcata gttccttact caaagccctc aacttatgtt gggggctttg tttatttttc
                                                                      4620
                                                                      4680
ttcggaaaat aagcgtgatc atgcgacagg tttgatgaca aaaaattagc gcaagaagac
aaaaatcacc ttgcgctaat gctctgttac aggtcactaa taccatctaa gtagttgatt
                                                                      4740
catagtgact gcatatgttg tgttttacag tattatgtag tctgtttttt atgcaaaatc
                                                                      4800
taatttaata tattgatatt tatatcattt tacgtttctc gttcagcttt tttatactaa
                                                                      4860
qttqqcatta taaaaaaqca ttqcttatca atttqttqca acqaacaqqt cactatcaqt
                                                                      4920
                                                                      4980
caaaataaaa tcattatttg atttcaattt tgtcccactc cctgcctctg tcatcacgat
actgtgatgc catggtgtcc gacttatgcc cgagaagatg ttgagcaaac ttatcgctta
                                                                      5040
totgettete ataqaqtett gcagacaaac tgegcaacte gtgaaaggta ggegceeeeg
                                                                      5100
gtcgaatttg ctttcgaatt tctgccattc atccgcttat tatcacttat tcaggcqtag
                                                                      5160
caccaggogt ttaagggcac caataactgc cttaaaaaaa ttacgccccg ccctgccact
                                                                      5220
catcgcagta ctgttgtaat tcattaagca ttctgccgac atggaagcca tcacagacgg
                                                                      5280
catgatgaac ctgaatcqcc agcggcatca gcaccttgtc gccttgcgta taatatttgc
                                                                      5340
                                                                      5400
ccatqqtqaa aacqqqqqcq aaqaaqttqt ccatattqqc cacqtttaaa tcaaaactqq
tgaaactcac ccagggattg gctgagacga aaaacatatt ctcaataaac cctttaggga
                                                                      5460
aataggccag gttttcaccg taacacgcca catcttgcga atatatgtgt agaaactgcc
                                                                      5520
ggaaatcgtc gtggtattca ctccagagcg atgaaaacgt ttcagtttgc tcatggaaaa
                                                                      5580
cggtgtaaca agggtgaaca ctatcccata tcaccagctc accgtctttc attgccatac
                                                                      5640
                                                                      5641
<210> 14
<211> 5670
<212> DNA
<213> Artificial Sequence
<220>
<223> pJMF4
<400> 14
gaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg gataaaactt
                                                                        60
gtgcttattt ttctttacgg tctttaaaaa ggccgtaata tccagctgaa cggtctggtt
                                                                       120
                                                                       180
ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat gccattggga
                                                                       240
tatatcaacg gtggtatatc cagtgatttt tttctccatt ttagcttcct tagctcctga
aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat ggtgaaagtt
                                                                       300
ggaacctctt acgtgccgat caacgtctca ttttcgccaa aagttggccc agggcttccc
                                                                       360
ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat
                                                                       420
                                                                       480
ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt
gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt tcagctactg
                                                                       540
acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtgcggc
                                                                       600
cgcactggct tactatgttg gcactgatga gggtgtcagt gaagtgcttc atgtggcagg
                                                                       660
                                                                       720
agaaaaaagg ctgcaccggt gcgtcagcag aatatgtgat acaggatata ttccgcttcc
                                                                       780
togotoactg actogotacg ctoggtogtt cgactgoggc gagoggaaat ggottacgaa
cggggcggag atttcctgga agatgccagg aagatactta acagggaagt gagagggccg
                                                                       840
                                                                       900
cggcaaagcc gtttttccat aggctccgcc cccctgacaa gcatcacgaa atctgacgct
caaatcaqtq qtgqcgaaac ccgacaggac tataaagata ccaggcgttt cccctggcgg
                                                                       960
ctccctcgtg cgctctcctg ttcctgcctt tcggtttacc ggtgtcattc cgctgttatg
                                                                      1020
geogegittg teteatteca egectgacae teagiteegg gtaggeagit egeteeaage
                                                                      1080
tggactgtat gcacgaaccc cccgttcagt ccgaccgctg cgccttatcc ggtaactatc
                                                                      1140
gtcttgagtc caacccggaa agacatgcaa aagcaccact ggcagcagcc actggtaatt
                                                                      1200
gatttagagg agttagtctt gaagtcatgc gccggttaag gctaaactga aaggacaagt
                                                                      1260
                                                                      1320
tttggtgact gcgctcctcc aagccagtta cctcggttca aagagttggt agctcagaga
                                                                      1380
accttcgaaa aaccgccctg caaggcggtt ttttcgtttt cagagcaaga gattacgcgc
agaccaaaac gatctcaaga agatcatctt atgcggccgc atcagataaa atatttctag
                                                                      1440
atgccqaact cagaaqtgaa acgccgtagc gccgatggta gtgtggggtc tccccatgcg
                                                                      1500
agagtaggga actgccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt
                                                                      1560
togttttatc tgttgtttgt cggtgaacgc tctcctgagt aggacaaatc cgccgggagc
                                                                      1620
                                                                      1680
ggatttgaac gttgcgaagc aacggcccgg agggtggcgg gcaggacgcc cgccataaac
tgccaggcat caaattaagc agaaggccat cctgacggat ggcctttttg cgtttctaca
                                                                      1740
```

aactcttcct	gtcgtcatat	ctacaagcca	tcccccgcat	gcattaatgt	gagttagctc	1800
actcattagg	caccccaggc	tttacacttt	atgcttccgg	ctcgtatgtt	gtgtggaatt	1860
gtgagcggat	aacaatttca	cacaggaaac	agctatgacc	atgattacgc	caagcttgca	1920
tacctacaga	tcgactctag	aggatccccg	ggtaccgage	tctgagcaac	aagttgtaaa	1980
ctcactooct	gataaccttg	atatccaata	tgaagtgtta	actaaccato	gtgctaacga	2040
agetettaca	tgccaagata	tagacacaga	atogocttct	totaacaaag	taaacatgac	2100
agguettgeg	caaggtgaag	ctattanata	222292556	cotatttact	tecacagest	2160
gettgttaac	caayycyaay	cigityacic	addayactyy	tetestates	cccacagcat	2220
tegtetgatt	ctggatgttg	acaacyayca	gricaaaarc	ccccgcgcaa	cygytyactt	
acataagcta	gaaccaacag	ataagtttga	eggetteget	geeggtgaag	agginging	2280
tccattggtt	ggtgaatact	ggcaactatt	tgaaactgac	ttcatgccgg	gtgcattcgt	2340
ttctgctcca	aacgcagaac	ctaagatgat	tgcttctcta	aatactgaag	atgttgcgtc	2400
ttttgtgacg	ggtcttgaag	gtaacaacct	aaaacgtaca	ccagatgaca	acaatgtatt	2460
tgcaaacgct	gtgtctcgtt	ttgagaaaaa	cgaagaccta	gcaacacaag	acgtatcaac	2520
cacgttacta	ccaacaccaa	tgcacgttga	agcgggtaaa	ggcaaagtag	atatcgcgga	2580
taatattaca	ctgcctaaag	acqcattcga	tgcgactcag	ttcgcagcga	ttcaagatcg	2640
tocagaagtg	gtaggtgtgg	acqttcqtqq	tgatcttcct	gtaagcatca	ctqttqttcc	2700
tocagacttc	accggtgaat	tagcaaaatc	togtocttac	gaaatgagca	tcaaaggcga	2760
cantattata	attaaagcgt	togaccaage	aggcgctttc	tacqcaqtac	aatctatctt	2820
taacetaata	gatagccaaa	atoctoatto	totaccacaa	ctotctatta	aagatgcgcc	2880
tagttttgat	taccgtggtg	taataataa	tataactcat	aacttccact	ctaaggacgc	2940
actoottoo	acgctagacc	tarcarc	gtacaagatg	aacaaacttc	accttcacct	3000
aacccccgca	acgccagacc	atttagaaat	coccatata	cctgagetga	cagaagtggg	3060
aaccgatgat	gaaggctggc	gittagaaat	cccgggtctg	cttgagetga	ttaaatataa	3120
tgctaaccgt	tgtttcgata	cacaayayaa	aagetgtta	ctgcctcage	traggereegg	
tccaacgaca	gacaactttg	gctctggcta	cttcagcaaa	geagactacg	tggaaatett	3180
gaaatacgcg	aaagcacgta	acattgaagt	gattccagaa	atcgatatgc	cagctcacgc	3240
tcgtgcagca	gtagtatcaa	tggaagctcg	ttacgaccgc	ctaatggaag	aaggtaaaga	3300
agctgaagcg	aacgaatacc	gtctgatgga	tcctcaagat	acatcaaacg	taacgacggt	3360
tcagttctac	aataagcaaa	gcttcatcaa	cccatgtatg	gaatcttcaa	ctcgctttgt	3420
tgataaggtg	atttcagaag	tggcagcaat	gcaccaagaa	gctggcgctc	cactaacaac	3480
ttggcacttc	ggtggtgacg	aagcgaagaa	catcaagcta	ggtgctggtt	tccaagacgt	3540
taacqcaqaa	gataaagtaa	gctggaaagg	cacgattgac	ctgtctaaac	aagacaagcc	3600
gtttgcacag	tctccacaat	gtcagacgct	aatcacagat	ggcacagtca	gtgactttgc	3660
teacetacea	agccacttcg	cagaagaagt	gtcgaagatt	gttgctgaga	aaggcattcc	3720
aaacttccaa	gcttggcaag	atggtttgaa	atacagtgac	ggcgaaaaag	cgttcgctac	3780
agaaaatact	cgcgtaaact	tctgggacgt	tctgtactgg	ggcggtactt	cctcagtgta	3840
casataatet	aagaaaggtt	acqacqtqat	totttctaac	ccagattacg	tgtacatgga	3900
tatecestae	gaagttgacc	cassacaca	toottactac	toogcaacac	gtgcaacgga	3960
tacyccatac	atgtttggct	ttacaccaca	gaacatgcct	caaaacacaa	aaacttctot	4020
Lacicglaag	atguttggct	ttgcaccaga	agetassata	gaagagaaa	ctttctacca	4080
agatcgcgat	ggcaatggct	ttactggtaa	aggigaaacc	gaagcgaaac	acticctacgg	4140
tctatctgca	caactttggt	ctgagacagt	acycaacyac	gagcaacacg	agracatggt	
attccctcgc	gtcctcgctg	ctgctcagcg	tgcatggcac	egtgetgaet	gggaaaacga	4200
ctacaaagtt	ggtgttgagt	actcgcaaaa	ctctaatcta	gttgataaag	categetaaa	4260
ccaagactac	aaccgctttg	cgaacgtact	tggtcaacgt	gaactggcta	agctagaaaa	4320
atcaggtatt	gactaccgcc	taccagtacc	aggtgcaaaa	gtagaagatg	gtaagctagc	4380
aatgaacgtt	cagttccctg	gcgtaacgct	tcaatactct	ctggatggtg	agaactggtt	4440
gacttatgca	gacaacgctc	gtccaaatgt	aactggtgaa	gtcttcatcc	gctcggtatc	4500
tgcgacaggt	gagaaggtaa	gccgtatcac	tagcgtgaaa	taatagcgct	cagtattcac	4560
taaaatcata	gttccttact	caaagccctc	aacttatgtt	gggggctttg	tttattttc	4620
ttcggaaaat	aagcgtgatc	ccccqqqqqc	gcctaccttt	cacgagttgc	gcagtttgtc	4680
tocaagactc	tatgagaagc	agataagcga	taagtttgct	caacatcttc	tcqqqcataa	4740
gtcggacacc	atggcatcac	agtatcgtga	tgacagagge	agggagtggg	acaaaattga	4800
aatcaaataa	tgattttatt	ttgactgata	gtgacctgtt	cattacaaca	aattgataag	4860
gastgetttt	ttataatgcc	aacttantat	aaaaaaaacta	ааспапааас	gtaaaatgat	4920
stantet	atatattaaa	ttagatttta	catagees	acactacata	atactotaaa	4980
acadatatca	atatattada	ntagatitity	tacttacate	atattactac	cototaaaaa	5040
acacaacata	tgcagtcact	atgaatcaac	tacttagatg	grarragiga	anagetete-	5100
agcattagcg	caaggtgatt	tttgtcttct	egegetaatt	ttttgtcatc	adacctgtcg	
catgatcatg	gggctgcagg	aattcgatgg	tcgaatttgc	tttcgaattt	ctgccattca	5160
tccgcttatt	atcacttatt	caggcgtagc	accaggcgtt	taagggcacc	aataactgcc	5220
ttaaaaaaat	tacgccccgc	cctgccactc	atcgcagtac	tgttgtaatt	cattaagcat	5280
tctgccgaca	tggaagccat	cacagacggc	atgatgaacc	tgaatcgcca	gcggcatcag	5340
caccttgtcg	ccttgcgtat	aatatttgcc	catggtgaaa	acgggggcga	agaagttgtc	5400
_						

```
catattggcc acgtttaaat caaaactggt gaaactcacc cagggattgg ctgagacgaa
                                                                      5460
aaacatattc tcaataaacc ctttagggaa ataggccagg ttttcaccgt aacacgccac
                                                                      5520
atcttgcgaa tatatgtgta gaaactgccg gaaatcgtcg tggtattcac tccagagcga
                                                                      5580
                                                                      5640
tgaaaacqtt tcaqtttgct catggaaaac ggtgtaacaa gggtgaacac tatcccatat
                                                                      5670
caccagetea eegtettea ttgccatacg
<210> 15
<211> 238
<212> DNA
<213> Artificial Sequence
<220>
<223> LacZ(/chitobiase Fusion
<220>
<221> CDS
<222> (119)...(240)
<221> -35 signal
<222> (44)...(50)
<223> Lac promoter
<221> -10_signal
<222> (69)...(74)
<223> Lac promoter
<221> protein bind
<222> (81)...(107)
<223> Lac repressor binding site
<221> protein_bind
<222> (7)...(34)
<223> CAP-cAMP binding site
<400> 15
gcatgcatta atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt
                                                                       60
ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagct
                                                                      118
atg acc atg att acg cca agc ttg cat gcc tgc agg tcg act cta gag
                                                                      166
Met Thr Met Ile Thr Pro Ser Leu His Ala Cys Arg Ser Thr Leu Glu
1
gat ccc cgg gta ccg agc tct gag caa caa gtt gta aac tca ctg gct
                                                                      214
Asp Pro Arg Val Pro Ser Ser Glu Gln Gln Val Val Asn Ser Leu Ala
                                 25
                                                                      238
gat aac ctt gat atc caa tat gaa
Asp Asn Leu Asp Ile Gln Tyr Glu
         35
<210> 16
<211> 40
<212> PRT
<213> Artificial Sequence
<223> LacZ(/chitobiase Fusion
<400> 16
Met Thr Met Ile Thr Pro Ser Leu His Ala Cys Arg Ser Thr Leu Glu
                                    10
                 5
Asp Pro Arg Val Pro Ser Ser Glu Gln Gln Val Val Asn Ser Leu Ala
                                 25
            20
Asp Asn Leu Asp Ile Gln Tyr Glu
```

40

35

<210> 17 <211> 22 <212> PRT <213> Artificial Sequence <223> dnaA/chitobiase Fusion <400> 17 Met Ser Leu Ser Leu Trp Gln Gln Cys Leu Ala Arg Leu Gln Asp Glu 5 Leu Val Pro Ser Ser Glu 20 <210> 18 <211> 2643 <212> DNA <213> Artificial Sequence <220> <221> CDS <222> (1)...(2643) <223> dnaA/Chitobiase Fusion <400> 18 gtg tca ctt tcg ctt tgg cag cag tgt ctt gcc cga ttg cag gat qag 48 Met Ser Leu Ser Leu Trp Gln Gln Cys Leu Ala Arg Leu Gln Asp Glu 10 ttg gta ccg agc tct gag caa caa gtt gta aac tca ctg gct gat aac 96 Leu Val Pro Ser Ser Glu Gln Gln Val Val Asn Ser Leu Ala Asp Asn ctt gat atc caa tat gaa gtg tta act aac cat ggt gct aac gaa ggt 144 Leu Asp Ile Gln Tyr Glu Val Leu Thr Asn His Gly Ala Asn Glu Gly ctt gcg tgc caa gat atg ggc gca gaa tgg gct tct tgt aac aaa gta 192 Leu Ala Cys Gln Asp Met Gly Ala Glu Trp Ala Ser Cys Asn Lys Val aac atg acg ctt gtt aac caa ggt gaa gct gtt gac tca aaa gat tgg 240 Asn Met Thr Leu Val Asn Gln Gly Glu Ala Val Asp Ser Lys Asp Trp gct att tac ttc cac agc att cgt ctg att ctg gat gtt gac aac gag 288 Ala Ile Tyr Phe His Ser Ile Arg Leu Ile Leu Asp Val Asp Asn Glu cag ttc aaa atc tct cgt gta acg ggt gac cta cat aag cta gaa cca 336 Gln Phe Lys Ile Ser Arg Val Thr Gly Asp Leu His Lys Leu Glu Pro 105 aca gat aag ttt gac ggc ttc gct gcc ggt gaa gag gtt gtt ctt cca 384 Thr Asp Lys Phe Asp Gly Phe Ala Ala Gly Glu Glu Val Val Leu Pro ttg gtt ggt gaa tac tgg caa cta ttt gaa act gac ttc atg ccg ggt 432 Leu Val Gly Glu Tyr Trp Gln Leu Phe Glu Thr Asp Phe Met Pro Gly

	130					135					140					
		gtt Val														480
		gaa Glu														528
cta Leu	aaa Lys	cgt Arg	aca Thr 180	cca Pro	gat Asp	gac Asp	aac Asn	aat Asn 185	gta Val	ttt Phe	gca Ala	aac Asn	gct Ala 190	gtg Val	tct Ser	576
cgt Arg	ttt Phe	gag Glu 195	aaa Lys	aac Asn	gaa Glu	gac Asp	cta Leu 200	gca Ala	aca Thr	caa Gln	gac Asp	gta Val 205	tca Ser	acc Thr	acg Thr	624
tta Leu	cta Leu 210	cca Pro	aca Thr	cca Pro	atg Met	cac His 215	gtt Val	gaa Glu	gcg Ala	ggt Gly	aaa Lys 220	ggc Gly	aaa Lys	gta Val	gat Asp	672
atc Ile 225	gcg Ala	gat Asp	ggt Gly	att Ile	gcg Ala 230	ctg Leu	cct Pro	aaa Lys	gac Asp	gca Ala 235	ttc Phe	gat Asp	gcg Ala	act Thr	cag Gln 240	720
		gcg Ala														768
ggt Gly	gat Asp	ctt Leu	cct Pro 260	gta Val	agc Ser	atc Ile	act Thr	gtt Val 265	gtt Val	cct Pro	gca Ala	gac Asp	ttc Phe 270	acc Thr	ggt Gly	816
gaa Glu	tta Leu	gca Ala 275	aaa Lys	tct Ser	ggt Gly	gct Ala	tac Tyr 280	gaa Glu	atg Met	agc Ser	atc Ile	aaa Lys 285	ggc Gly	gac Asp	ggt Gly	864
att Ile	gtg Val 290	att Ile	aaa Lys	gcg Ala	ttc Phe	gac Asp 295	caa Gln	gca Ala	ggc Gly	gct Ala	ttc Phe 300	tac Tyr	gca Ala	gta Val	caa Gln	912
tct Ser 305	atc Ile	ttt Phe	ggc Gly	ctg Leu	gta Val 310	gat Asp	agc Ser	caa Gln	aat Asn	gct Ala 315	gat Asp	tct Ser	cta Leu	cca Pro	caa Gln 320	960
ctg Leu	tct Ser	att Ile	aaa Lys	gat Asp 325	gcg Ala	cct Pro	cgt Arg	ttt Phe	gat Asp 330	tac Tyr	cgt Arg	ggt Gly	gtg Val	atg Met 335	gtg Val	1008
gat Asp	gtg Val	gct Ala	cgt Arg 340	aac Asn	ttc Phe	cac His	tċt Ser	aag Lys 345	gac Asp	gca Ala	atc Ile	ctt Leu	gca Ala 350	acg Thr	cta Leu	1056
gac Asp	caa Gln	atg Met 355	gca Ala	gcg Ala	tac Tyr	aag Lys	atg Met 360	aac Asn	aaa Lys	ctt Leu	cac His	ctt Leu 365	cac His	cta Leu	acc Thr	1104
gat Asp	gat Asp 370	gaa Glu	ggc Gly	tgg Trp	cgt Arg	tta Leu 375	gaa Glu	atc Ile	ccg Pro	ggt Gly	ctg Leu 380	cct Pro	gag Glu	ctg Leu	aca Thr	1152

gaa gt Glu Va 385															1200
ctg co Leu Pr	-						_		-						1248
tac tt Tyr Ph															1296
cgt aa Arg As		Glu													1344
gca gc Ala Al 45	a Val														1392
ggt aa Gly Ly 465															1440
aca to Thr Se															1488
aac co Asn Pr															1536
gaa gt Glu Va	g gca 1 Ala 515	gca Ala	atg Met	cac His	caa Gln	gaa Glu 520	gct Ala	ggc Gly	gct Ala	cca Pro	cta Leu 525	aca Thr	act Thr	tgg Trp	1584
cac tt His Ph 53	e Gly	ggt Gly	gac Asp	gaa Glu	gcg Ala 535	aag Lys	aac Asn	atc Ile	aag Lys	cta Leu 540	ggt Gly	gct Ala	ggt Gly	ttc Phe	1632
caa ga Gln As 545															1680
ctg to Leu Se	t aaa r Lys	caa Gln	gac Asp 565	aag Lys	ccg Pro	ttt Phe	gca Ala	cag Gln 570	tct Ser	cca Pro	caa Gln	tgt Cys	cag Gln 575	acg Thr	1728
cta at Leu Il	c aca e Thr	gat Asp 580	ggc Gly	aca Thr	gtc Val	agt Ser	gac Asp 585	ttt Phe	gct Ala	cac His	cta Leu	cca Pro 590	agc Ser	cac His	1776
ttc gc Phe Al	g gaa a Glu 595	gaa Glu	gtg Val	tcg Ser	aag Lys	att Ile 600	gtt Val	gct Ala	gag Glu	aaa Lys	ggc Gly 605	att Ile	cca Pro	aac Asn	1824
ttc ca Phe Gl 61	n Ala														1872

								aac Asn								1920
ggc Gly	ggt Gly	act Thr	tcc Ser	tca Ser 645	gtg Val	tac Tyr	gag Glu	tgg Trp	tct Ser 650	aag Lys	aaa Lys	ggt Gly	tac Tyr	gac Asp 655	gtg Val	1968
								tac Tyr 665								2016
gac Asp	ccg Pro	aaa Lys 675	Glu	cgt Arg	ggt Gly	tac Tyr	tac Tyr 680	tgg Trp	gca Ala	aca Thr	cgt Arg	gca Ala 685	acg Thr	gat Asp	act Thr	2064
								gag Glu								2112
								ggc Gly								2160
								tct Ser								2208
								tac Tyr 745								2256
								cgt Arg								2304
								aac Asn								2352
tcg Ser 785	cta Leu	aac Asn	caa Gln	Asp	tac Tyr 790	aac Asn	cgc Arg	ttt Phe	gcg Ala	aac Asn 795	gta Val	ctt Leu	ggt Gly	caa Gln	cgt Arg 800	2400
gaa Glu	ctg Leu	gct Ala	aag Lys	cta Leu 805	gaa Glu	aaa Lys	tca Ser	ggt Gly	att Ile 810	gac Asp	tac Tyr	cgc Arg	cta Leu	cca Pro 815	gta Val	2448
								aag Lys 825								2496
cct Pro	ggc Gly	gta Val 835	acg Thr	ctt Leu	caa Gln	tac Tyr	tct Ser 840	ctg Leu	gat Asp	ggt Gly	gag Glu	aac Asn 845	tgg Trp	ttg Leu	act Thr	2544
								gta Val								2592
tcg	gta	tct	gcg	aca	ggt	gag	aag	gta	agc	cgt	atc	act	agc	gtg	aaa	2640

Ser Val Ser Ala Thr Gly Glu Lys Val Ser Arg Ile Thr Ser Val Lys

870 taa 2643 <210> 19 <211> 880 <212> PRT <213> Artificial Sequence <220> <223> dnaA/Chitobiase Fusion <400> 19 Met Ser Leu Ser Leu Trp Gln Gln Cys Leu Ala Arg Leu Gln Asp Glu 10 Leu Val Pro Ser Ser Glu Gln Gln Val Val Asn Ser Leu Ala Asp Asn 25 Leu Asp Ile Gln Tyr Glu Val Leu Thr Asn His Gly Ala Asn Glu Gly 40 Leu Ala Cys Gln Asp Met Gly Ala Glu Trp Ala Ser Cys Asn Lys Val . 55 Asn Met Thr Leu Val Asn Gln Gly Glu Ala Val Asp Ser Lys Asp Trp 70 Ala Ile Tyr Phe His Ser Ile Arg Leu Ile Leu Asp Val Asp Asn Glu 85 90 Gln Phe Lys Ile Ser Arg Val Thr Gly Asp Leu His Lys Leu Glu Pro 100 105 Thr Asp Lys Phe Asp Gly Phe Ala Ala Gly Glu Glu Val Val Leu Pro 115 120 125 Leu Val Gly Glu Tyr Trp Gln Leu Phe Glu Thr Asp Phe Met Pro Gly 135 Ala Phe Val Ser Ala Pro Asn Ala Glu Pro Lys Met Ile Ala Ser Leu 150 155 Asn Thr Glu Asp Val Ala Ser Phe Val Thr Gly Leu Glu Gly Asn Asn 165 170 Leu Lys Arg Thr Pro Asp Asp Asn Asn Val Phe Ala Asn Ala Val Ser 185 180 Arg Phe Glu Lys Asn Glu Asp Leu Ala Thr Gln Asp Val Ser Thr Thr 200 Leu Leu Pro Thr Pro Met His Val Glu Ala Gly Lys Gly Lys Val Asp 215 220 Ile Ala Asp Gly Ile Ala Leu Pro Lys Asp Ala Phe Asp Ala Thr Gln 230 235 Phe Ala Ala Ile Gln Asp Arg Ala Glu Val Val Gly Val Asp Val Arg 245 250 Gly Asp Leu Pro Val Ser Ile Thr Val Val Pro Ala Asp Phe Thr Gly 265 Glu Leu Ala Lys Ser Gly Ala Tyr Glu Met Ser Ile Lys Gly Asp Gly 280 285 Ile Val Ile Lys Ala Phe Asp Gln Ala Gly Ala Phe Tyr Ala Val Gln 295 300 Ser Ile Phe Gly Leu Val Asp Ser Gln Asn Ala Asp Ser Leu Pro Gln 310 315 Leu Ser Ile Lys Asp Ala Pro Arg Phe Asp Tyr Arg Gly Val Met Val 330 325 Asp Val Ala Arg Asn Phe His Ser Lys Asp Ala Ile Leu Ala Thr Leu 345 340 Asp Gln Met Ala Ala Tyr Lys Met Asn Lys Leu His Leu His Leu Thr 360

Asp Asp Glu Gly Trp Arg Leu Glu Ile Pro Gly Leu Pro Glu Leu Thr 370 375 Glu Val Gly Ala Asn Arg Cys Phe Asp Thr Gln Glu Lys Ser Cys Leu 390 Leu Pro Gln Leu Gly Ser Gly Pro Thr Thr Asp Asn Phe Gly Ser Gly 405 410 Tyr Phe Ser Lys Ala Asp Tyr Val Glu Ile Leu Lys Tyr Ala Lys Ala 420 425 Arg Asn Ile Glu Val Ile Pro Glu Ile Asp Met Pro Ala His Ala Arg 440 Ala Ala Val Val Ser Met Glu Ala Arg Tyr Asp Arg Leu Met Glu Glu 455 460 Gly Lys Glu Ala Glu Ala Asn Glu Tyr Arg Leu Met Asp Pro Gln Asp 470 475 Thr Ser Asn Val Thr Thr Val Gln Phe Tyr Asn Lys Gln Ser Phe Ile 485 490 Asn Pro Cys Met Glu Ser Ser Thr Arg Phe Val Asp Lys Val Ile Ser 505 500 Glu Val Ala Ala Met His Gln Glu Ala Gly Ala Pro Leu Thr Thr Trp 520 His Phe Gly Gly Asp Glu Ala Lys Asn Ile Lys Leu Gly Ala Gly Phe 535 Gln Asp Val Asn Ala Glu Asp Lys Val Ser Trp Lys Gly Thr Ile Asp 550 555 Leu Ser Lys Gln Asp Lys Pro Phe Ala Gln Ser Pro Gln Cys Gln Thr 565 570 Leu Ile Thr Asp Gly Thr Val Ser Asp Phe Ala His Leu Pro Ser His 585 Phe Ala Glu Glu Val Ser Lys Ile Val Ala Glu Lys Gly Ile Pro Asn 600 Phe Gln Ala Trp Gln Asp Gly Leu Lys Tyr Ser Asp Gly Glu Lys Ala 615 620 Phe Ala Thr Glu Asn Thr Arg Val Asn Phe Trp Asp Val Leu Tyr Trp 630 635 Gly Gly Thr Ser Ser Val Tyr Glu Trp Ser Lys Lys Gly Tyr Asp Val 645 650 Ile Val Ser Asn Pro Asp Tyr Val Tyr Met Asp Met Pro Tyr Glu Val 665 670 Asp Pro Lys Glu Arg Gly Tyr Tyr Trp Ala Thr Arg Ala Thr Asp Thr 675 680 Arg Lys Met Phe Gly Phe Ala Pro Glu Asn Met Pro Gln Asn Ala Glu 695 Thr Ser Val Asp Arg Asp Gly Asn Gly Phe Thr Gly Lys Gly Glu Ile 710 715 Glu Ala Lys Pro Phe Tyr Gly Leu Ser Ala Gln Leu Trp Ser Glu Thr 725 730 Val Arg Asn Asp Glu Gln Tyr Glu Tyr Met Val Phe Pro Arg Val Leu 745 . Ala Ala Ala Gln Arg Ala Trp His Arg Ala Asp Trp Glu Asn Asp Tyr 760 Lys Val Gly Val Glu Tyr Ser Gln Asn Ser Asn Leu Val Asp Lys Ala 775 780 Ser Leu Asn Gln Asp Tyr Asn Arg Phe Ala Asn Val Leu Gly Gln Arg 790 795 Glu Leu Ala Lys Leu Glu Lys Ser Gly Ile Asp Tyr Arg Leu Pro Val 805 810 Pro Gly Ala Lys Val Glu Asp Gly Lys Leu Ala Met Asn Val Gln Phe 825 Pro Gly Val Thr Leu Gln Tyr Ser Leu Asp Gly Glu Asn Trp Leu Thr 840 845 Tyr Ala Asp Asn Ala Arg Pro Asn Val Thr Gly Glu Val Phe Ile Arg

850 855 860 Ser Val Ser Ala Thr Gly Glu Lys Val Ser Arg Ile Thr Ser Val Lys 865 870 875 880

-17-

U.S.S.N.: 10/521,300

Filed: 1/14/05

Case No.: 21126YP

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



- 1 (CLD 2000) (CD200) (CD200) (CD300) (CD300) (CD300) (CD300) (CD300) (CD300) (CD300) (CD300) (CD300) (CD300)

(43) International Publication Date 19 April 2001 (19.04.2001)

PCT

(10) International Publication Number WO 01/27322 A3

(51) International Patent Classification7:

101

- (21) International Application Number: PCT/US00/21048
- (22) International Filing Date: 2 August 2000 (02.08.2000)
- (25) Filing Language:

English

C12Q 1/68

(26) Publication Language:

English

(30) Priority Data:

60/159,221

13 October 1999 (13.10.1999) US

- (71) Applicant (for all designated States except US): ELITRA PHARMACEUTICALS, INC. [US/US]; Suite A, 3510 Dunhill Street, San Diego, CA 92121 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): ZYSKIND, Judith [US/US]; 8514 La Jolla Scenic Drive, La Jolla, CA 92047 (US).
- (74) Agent: HUNT, Dale, C.; 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

- (81) Designated States (national): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KR (utility model), KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

with international search report

(88) Date of publication of the international search report: 13 December 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

7322 A3

(54) Title: CHITOBIASE AS A REPORTER ENZYME

(57) Abstract: The present invention relates to reporter gene constructs encoding a cytoplasmic form of chitobiase (N,N'-diacetyl-chitobiase) and methods of using these reporter gene constructs. The use of a cytoplasmic form of chitobiase as a reporter enzyme is generally applicable to the study of gene expression in organisms which do not contain N-acetyl-β-D-glucosaminidases.

INTERNATIONAL SEARCH REPORT

International Application No Pt. i/US 00/21048

	SIFICATION OF SUBJECT MATTER C12Q1/68									
	to International Patent Classification (IPC) or to both national classificatio	fication and IPC	· · · · · · · · · · · · · · · · · · ·							
Minimum do	ocumentation searched (classification system tollowed by classifica-	ation symbols)								
IPC 7										
Documentat	Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched									
1	data base consulted during the international search (name of data b	base and, where practical, search terms used	1)							
EPO-In	ternal, WPI Data, BIOSIS									
	ENTS CONSIDERED TO BE RELEVANT									
Category °	Citation of document, with indication, where appropriate, of the n	elevant passages	Relevant to claim No.							
X	KALABAT D Y ET AL: "Chitobiase, reporter enzyme" BIOTECHNIQUES,US,EATON PUBLISHIN vol. 25, no. 6, December 1998 (1 pages 1030-1035, XP002106689 ISSN: 0736-6205 cited in the application the whole document	NG, NATICK,	1-50							
А	WO 98 49320 A (AHLUWALIA NAVNEET BERNHARD F (CA); FALCONER MARCIA 5 November 1998 (1998-11-05) abstract page 1 -page 2		1–50							
X Furth	ner documents are listed in the continuation of box C.	Patent family members are listed in	in annex.							
Special cate	tegories of cited documents :	"T" later document published after the inter-								
"A" documer conside	document defining the general state of the art which is not cited to understand the principle or theory underlying the									
	locument but published on or after the international	invention "X" document of particular relevance; the classification of the considered payel or cannot be								
"L" documen which is citation	cannot be considered novel or cannot be considered to to document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another clation or other special reason (as specified) cannot be considered novel or cannot be considered n									
other men	nnt referring to an oral disclosure, use, exhibition or neans nt published prior to the International filing date but an the priority date claimed	document is combined with one or mon	document is combined with one or more other such docu- ments, such combination being obvious to a person skilled in the art.							
	actual completion of the international search	Date of mailing of the international sear								
	July 2001	23/07/2001								
Name and m	nailing address of the ISA	Authorized officer								
	European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3018	Hart-Davis, J								

1

INTERNATIONAL SEARCH REPORT

International Application No
Pull/US 00/21048

		Pui/US 00/21048
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; March 1998 (1998-03) PIRUZIAN E S ET AL: "The use of a thermostable beta-glucanase gene from Clostridium thermocellum as a reporter gene in plants." Database accession no. PREV199800227889 XP002171323 abstract & MOLECULAR & GENERAL GENETICS, vol. 257, no. 5, March 1998 (1998-03), pages 561-567, ISSN: 0026-8925	1-50

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intrational Application No

		*****		00/21048
Patent document cited in search report	Publication date	F	Patent family member(s)	Publication date
WO 9849320 A	05-11-1998	CA AU	2203613 A 6085698 A	24-10-1998 24-11-1998

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
☐ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
□ OTHER:

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)